

ANNALES DE L'I. H. P.

PAUL BÉTHOUX

**Tests et estimations concernant certaines fonctions
aléatoires, en particulier laplaciennes**

Annales de l'I. H. P., tome 17, n° 4 (1962), p. 255-322

http://www.numdam.org/item?id=AIHP_1962__17_4_255_0

© Gauthier-Villars, 1962, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P. » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

Tests et estimations concernant certaines fonctions aléatoires, en particulier laplaciennes

par

Paul BÉTHOUX.

SOMMAIRE.

Étude de quelques problèmes de statistique, à partir de l'enregistrement continu d'une réalisation d'une fonction aléatoire $X(t)$, sur un intervalle de temps fini J : Étude de la vitesse de transmission de l'information à travers un canal de communication de bande passante finie. Estimation par le maximum de vraisemblance de la moyenne $\varphi(t) = E[X(t)]$ de la fonction aléatoire $X(t)$, prévision linéaire lorsque les lois temporelles des fonctions aléatoires considérées contiennent des paramètres inconnus.

Introduction. — Dans le chapitre I, nous considérons le test suivant sur la moyenne d'une fonction aléatoire; $X(t)$ est une fonction aléatoire, permanente, laplacienne, réelle; sa moyenne $E[X(t)]$ appartient à un ensemble de n fonctions qui sont les transformées de Fourier de fonctions ayant pour support un intervalle $(-\Omega, +\Omega)$; sa partie centrée $U(t)$ est une fonction aléatoire, stationnaire, de spectre connu et ayant le même support $(-\Omega, +\Omega)$; nous considérons à partir de l'enregistrement d'une réalisation de $X(t)$, sur un intervalle fini J , le problème de la discrimination entre les n possibilités pour la moyenne $E[X(t)]$.

En remarquant que $X(t)$ pour $t \in J$ peut être considérée comme un élément aléatoire X à valeurs dans \mathcal{X} , espace de Hilbert des fonctions réelles, définies sur J et de carré du module sommable par rapport à la mesure de Lebesgue sur cet intervalle, nous avons à tester entre n lois possibles et données pour X . Alors, comme il a été indiqué par M. Fortet

et M^{lle} Mourier, nous pouvons utiliser comme résumé exhaustif les rapports de vraisemblance.

Dans le cas $n = 2$, nous calculons l'erreur de discrimination en fonction des données et dans le cas général nous montrons comment le test est équivalent à un test plus simple sur les moyennes de $(n - 1)$ variables aléatoires laplaciennes, indépendantes.

Nous remarquons que, par suite de l'analyticité des trajectoires, l'erreur de discrimination est indépendante de la longueur de l'intervalle d'observation J : ceci nous montre qu'un modèle supposant un spectre strictement limité à une bande finie, bien que couramment introduit par les ingénieurs ne reflète pas exactement la réalité physique.

Nous appliquons néanmoins les résultats obtenus au problème de l'étude de la vitesse de transmission de l'information dans un canal de bande passante finie, perturbée par un bruit blanc, le temps est réintroduit ici de façon, semble-t-il, naturelle en supposant que l'émetteur fonctionnant pendant une durée T ne peut émettre que des signaux d'énergie totale inférieure ou égale à PT . Nous obtenons certaines formules de majoration et de minoration. Les résultats obtenus sont incompatibles avec la formule avancée de Shannon pour la capacité d'un tel canal.

Dans le chapitre II, nous cherchons à estimer par la règle du maximum de vraisemblance, la valeur moyenne $E[X(t)]$ d'une fonction aléatoire, réelle, laplacienne, définie sur un certain intervalle fini J , de covariance continue et connue, dont nous observons sur J une réalisation.

Ce problème a été abordé antérieurement et simultanément par plusieurs autres auteurs (Mann, Striabel, Parzen, etc) soit pour des processus particuliers (Wiener-Lévy, Ornstein-Uhlenbeck) soit pour une fonction aléatoire laplacienne ne vérifiant que les hypothèses mentionnées plus haut, mais toujours dans le cas où l'on suppose que *a priori* $E[X(t)]$ est une combinaison linéaire inconnue de fonctions connues, définies sur J . Dans ce cas, le problème se réduit à une estimation d'un nombre fini de scalaires réels; les coefficients de la combinaison linéaire.

Nous avons essayé ici de poser le problème de façon plus générale.

Si nous supposons que *a priori* $E[X(t)] \in \mathfrak{X}$, alors $X(t)$, peut être considérée comme un élément aléatoire à valeurs dans \mathfrak{X} . Si nous désignons par μ_ρ la loi de X sous l'hypothèse $E[X(t)] \equiv \rho(t)$, alors sur

l'espace H des ρ tels que μ_ρ est absolument continue par rapport à μ_0 , on peut définir un produit scalaire avec lequel H est un espace de Hilbert.

Nous supposons que *a priori* $E[X(t)]$ appartient à un certain sous-ensemble \mathcal{F} de H et nous définissons ce que nous entendons par estimateur du maximum de vraisemblance. Dans le cas où \mathcal{F} est un sous-espace vectoriel fermé de H , nous avons établi un théorème qui montre la difficulté, sinon l'impossibilité, de définir un estimateur du maximum de vraisemblance dans le cas où \mathcal{F} est de dimension infinie : si \mathcal{E} est l'ensemble des $x \in \mathcal{X}$ pour lesquels la série qui définit le logarithme de la densité de Radon-Nikodym de μ_ρ par rapport à μ_0 , converge pour tout $\rho \in \mathcal{F}$, alors $\mu_0(\mathcal{E})=1$ si \mathcal{F} est de dimension finie et $\mu_0(\mathcal{E})=0$ si \mathcal{F} est de dimension infinie. Dans le cas où \mathcal{F} est un sous-espace vectoriel de dimension finie de H , nous étudions en détail toutes les propriétés de l'estimateur $\hat{\rho}$ du maximum de vraisemblance qui se trouve être non biaisé, exhaustif, à variance minimale. De plus nous remarquons que si l'on abandonne le caractère laplacien de $X(t)$, l'estimateur $\hat{\rho}$ a les mêmes propriétés d'optimalité mais cette fois seulement parmi les estimateurs linéaires, non biaisés.

Nous examinons ensuite trois autres cas où nous avons pu définir un estimateur du maximum de vraisemblance :

1° \mathcal{F} est un sous-ensemble convexe et fermé d'un sous-espace vectoriel de dimension finie de H .

2° \mathcal{F} est un sous-ensemble convexe et compact de $\Gamma(\mathcal{X})$ où Γ désigne l'opérateur linéaire dans \mathcal{X} engendré par le noyau $\Gamma(t, \tau)$, covariance de $X(t)$.

3° \mathcal{F} est l'ensemble des fonctions dépendant du paramètre τ ,

$$\rho_\tau(t), \quad \text{où} \quad \rho_\tau(t) = \rho_0(t - \tau) \quad \text{et} \quad \Gamma(t, \tau) = r(t - \tau).$$

Le chapitre III est consacré à certains problèmes de prévision à partir de l'observation d'une fonction aléatoire permanente sur un intervalle fini. Supposons que nous observons une fonction aléatoire $X_t(\omega)$ sur l'intervalle J et désignons par \mathcal{B}_J le plus petit σ -corps par rapport auquel les $X_t(\omega)$ sont mesurables pour tout $t \in J$. Parmi toutes les fonctions numériques mesurables par rapport à \mathcal{B}_J , la meilleure prévision au sens des moindres carrés d'un nombre aléatoire $Y(\omega)$, plus ou moins corrélé

avec la fonction aléatoire $X_t(\omega)$ est l'espérance mathématique conditionnelle de $Y(\omega)$ par rapport à \mathcal{B}_t . Dans le cas où $X_t(\omega)$ est une fonction aléatoire continue en probabilité, alors la meilleure prévision est la limite de la meilleure prévision à partir de l'observation de $X_t(\omega)$ en un nombre fini de points de J , quand le maximum de la distance entre deux consécutifs de ces points tend vers zéro. Peut être ces résultats (au moins en ce qui concerne le premier) sont-ils connus, mais à ma connaissance ils ne figurent explicitement nulle part.

Nous étudions ensuite un problème particulier de filtrage, intéressant la théorie de l'information, qui donne un exemple où la prévision optimale (au sens des moindres carrés) n'est pas forcément linéaire et qui fournit également une application des densités de probabilités sur \mathcal{X} introduite aux chapitres précédents.

Enfin, nous considérons un problème de prévision linéaire où nous avons affaire à des fonctions aléatoires dont la moyenne est une combinaison linéaire inconnue de fonctions données. Plus précisément $X(t)$ et $Y(t)$ désignent deux fonctions aléatoires du second ordre, de la forme

$$X(t) = m(t) + X_0(t), \quad Y(t) = m(t) + Y_0(t), \quad \text{avec} \quad E[X_0(t)] \equiv E[Y_0(t)] \equiv 0$$

et

$$m(t) = \sum_{k=1}^n a_k g_k(t)$$

où les a_k sont des nombres certains inconnus et où les $g_k(t)$ sont des fonctions connues. Nous caractérisons la meilleure prévision linéaire, non biaisée $\tilde{Y}(\tau)$ de $Y(\tau)$ ($\tau > 0$) à partir de l'observation de $X(t)$ pour $t \in [-T, 0]$. Puis nous montrons comment on peut obtenir explicitement $\tilde{Y}(\tau)$ par les méthodes introduites par Wiener et Yaglom, dans le cas où $X_0(t)$ et $Y_0(t)$ sont stationnaires et stationnairement corrélés et admettent des densités spectrales de corrélation et d'intercorrélation rationnelles et pour une certaine classe de fonctions $g_k(t)$ [englobant le cas où $m(t)$ est un polynôme et celui où $m(t)$ est un polynôme trigonométrique].

Dans un article qui vient de paraître sur la prévision multidimensionnelle, Yaglom énonce des résultats semblables.

*
* *

Je suis heureux d'exprimer ici mes plus vifs remerciements à M. le Professeur FORTET pour avoir bien voulu constamment orienter et aider mes recherches.

Je lui suis infiniment reconnaissant pour ses multiples observations et conseils qui m'ont permis de faire progresser ce travail.

CHAPITRE I.

TESTS SUR LA MOYENNE D'UNE FONCTION ALÉATOIRE.

Application au problème de la vitesse de transmission de l'information à travers un canal de communication de bande passante finie.

1. Test sur la moyenne d'une fonction aléatoire. — Dans la pratique, en particulier en technique des communications, on rencontre souvent des problèmes du type suivant :

Étant donné une fonction aléatoire $X(t)$ définie sur un intervalle I , tester à partir de l'observation d'une réalisation de $X(t)$, c'est-à-dire une seule observation de toutes les valeurs de $X(t)$ pour $t \in I$, la moyenne $E[X(t)]$ de cette fonction aléatoire, sachant que cette moyenne est égale à l'une des fonctions $\rho_i(t)$, ($i = 1, \dots, n$). De tels tests ont été étudiés dans le cas général notamment par Grenander et Adhikari (voir $[G_1]$, $[A_1]$).

Nous considérerons ici, des tests de cette sorte, pour une fonction aléatoire $X(t)$, et des fonctions $\rho_k(t)$ particulières.

Remarquons que nous pouvons toujours supposer $I = [0, T]$ et écrire

$$(1.1) \quad X(t) = U(t) + \rho_k(t),$$

où $U(t)$ est une fonction aléatoire telle que $E[U(t)] = 0$ pour tout $t \in [0, T]$.

Nous supposons que $U(t)$ est une fonction aléatoire, réelle, mesurable, Laplacienne, de covariance $\Gamma(t, \tau)$ connue et continue.

Alors presque toute réalisation de $U(t)$ est de carré sommable sur $[0, T]$, par rapport à la mesure de Lebesgue sur cet intervalle.

Soit \mathcal{X} l'espace de Hilbert, séparable, des fonctions réelles $h(t)$ de la variable réelle t sur l'intervalle $[0, T]$ qui sont de carré sommable

sur cet intervalle par rapport à la mesure de Lebesgue. (Munissons \mathcal{X} du produit scalaire et de la topologie habituels).

$U(t)$ peut-être considéré comme un élément aléatoire U à valeurs dans \mathcal{X} . La donnée de la loi temporelle de $U(t)$ [déterminée en l'occurrence par $\Gamma(t, \tau)$] est équivalente à la connaissance d'une mesure μ_0 sur \mathcal{A} , plus petit corps de Borel de sous ensembles de \mathcal{X} , qui rend mesurables les applications

$$(1.2) \quad x \xrightarrow{\omega_t(\cdot)} x(t)$$

de $(\mathcal{X}, \mathcal{A})$ dans (\mathbb{R}, ζ) (ζ boréliens de la droite \mathbb{R}), pour tout $t \in [0, T]$.

μ_0 est de plus prolongeable de façon unique sur $\overline{\mathcal{A}}$ complétion de \mathcal{A} par rapport à μ_0 (voir [B₂]).

Si nous supposons, ce que nous ferons dans la suite, que :

$$(1.3) \quad \rho_i(t) \in \mathcal{X} \quad \text{pour } i = 1, \dots, n,$$

alors de même $V_i(t) = U(t) + \rho_i(t)$ peut être considéré comme un élément aléatoire à valeurs dans \mathcal{X} .

Soit μ_i la mesure définie sur \mathcal{A} dont la connaissance est équivalente à celle de la loi temporelle de $V_i(t)$ [déterminée par $\Gamma(t, \tau)$ et $\rho_i(t)$].

Nous pouvons donc considérer le problème du test entre les différents $\rho_i(t)$, $i = 1, \dots, n$, connaissant $\Gamma(t, \tau)$ comme le problème suivant :

Étant donné une observation de l'élément aléatoire X défini sur l'espace mesurable $(\mathcal{X}, \mathcal{A})$, tester à laquelle des lois μ_1, \dots, μ_n il obéit.

Désignons par $\overline{\mathcal{A}}^{(i)}$ la complétion de $\overline{\mathcal{A}}$ par rapport à μ_i . Alors les éléments aléatoires U et V_i respectivement définis par $(\mathcal{X}, \overline{\mathcal{A}}, \mu_0)$ et $(\mathcal{X}, \overline{\mathcal{A}}^{(i)}, \mu_i)$ sont des l -éléments aléatoires sur \mathcal{X} , au sens de M^{lle} Mourier (voir [M₂]).

Montrons-le, par exemple, pour U :

Soit h^* une fonctionnelle linéaire fortement continue sur \mathcal{X} ; on a

$$h^*(x) = \int_0^T x(t) h(t) dt, \quad h(t) \in \mathcal{X}.$$

Subdivisons l'intervalle $[0, T]$ par des points

$$0 = t_{1,m} < \dots < t_{\alpha,m} \dots < t_{m,m} = T$$

et soit :

$$h'_m(x) = \sum_{\alpha=1}^{m-1} \omega_{t_{\alpha,m}}(x) h(t_{\alpha,m}) [t_{\alpha+1,m} - t_{\alpha,m}].$$

Supposons que $\sup_{\alpha} |t_{\alpha+1, m} - t_{\alpha, m}| \rightarrow 0$, quand $m \rightarrow \infty$ $h'_m(x)$ est une fonction mesurable \mathcal{A} et $h'_m(x)$ converge quand $m \rightarrow \infty$ en μ_0 -mesure vers

$$h'(x) = \int_0^T \omega_t(x) h(t) dt,$$

intégrale en moyenne quadratique.

Donc il existe une sous-suite m_k telle que

$$h'_{m_k} \rightarrow h'(x) \quad \text{presque sûrement quand } k \rightarrow \infty.$$

Par suite, $h'(x)$ est une fonction $\overline{\mathcal{A}}$ -mesurable et comme $h'(x) = h^*(x)$ presque sûrement, il en est de même pour $h^*(x)$. Alors si nous désignons par \mathcal{B} le σ -corps engendré par les ensembles « cylindriques » définis par les fonctionnelles linéaires fortement continues sur \mathcal{X} , qui est aussi le σ -corps engendré par les sphères de \mathcal{X} , alors on a

$$\mathcal{B} \subset \overline{\mathcal{A}}, \quad \mathcal{B} \subset \overline{\mathcal{A}^{(i)}} \quad (i = 1, \dots, n).$$

Si les mesures $\mu_0, \mu_1, \dots, \mu_n$ sont toutes équivalentes, alors

$$\overline{\mathcal{B}} = \overline{\mathcal{A}^{(i)}} = \overline{\mathcal{A}} \quad \text{pour tout } i = 1, \dots, n$$

($\overline{\mathcal{B}}$ étant la complétion de \mathcal{B}).

En effet, $\frac{1}{\Delta t} \int_t^{t+\Delta t} \omega_{\tau}(x) d\tau$ est une application mesurable de $(\mathcal{X}, \mathcal{B})$ dans (\mathbb{R}, ζ) et

$$\begin{aligned} & \int_{\mathcal{X}} \left| \omega_t(x) - \frac{1}{\Delta t} \int_t^{t+\Delta t} \omega_{\tau}(x) d\tau \right|^2 d\mu_0(x) \\ &= \Gamma(t, t) + \frac{1}{(\Delta t)^2} \int_t^{t+\Delta t} \int_t^{t+\Delta t} \Gamma(\tau, \tau') d\tau d\tau' - \frac{2}{\Delta t} \int_t^{t+\Delta t} \Gamma(t, \tau) d\tau \\ & \rightarrow 0 \quad \text{quand } t \rightarrow 0. \end{aligned}$$

Donc, pour t fixé :

$$\omega_t(x) - m \int_t^{t+\frac{1}{m}} \omega_{\tau}(x) d\tau \rightarrow 0$$

en moyenne quadratique μ_0 , donc en mesure par rapport à μ_0 quand $m \rightarrow \infty$.

Il existe alors une sous-suite m_k telle que

$$\omega_l(x) - m_k \int_l^{l + \frac{1}{m_k}} \omega_\tau(x) d\tau \rightarrow 0 \quad \text{presque sûrement } \mu_0 \text{ quand } k \rightarrow \infty.$$

Et par suite $\omega_l(x)$ est une application mesurable de $(\mathcal{X}, \overline{\mathcal{B}})$ dans (\mathbb{R}, ζ) .

D'où

$$\mathcal{A} \subset \overline{\mathcal{B}}.$$

Et comme on a vu tout à l'heure que $\mathcal{B} \subset \overline{\mathcal{A}}$, on a bien

$$\overline{\mathcal{B}} = \overline{\mathcal{A}} = \overline{\mathcal{A}}^{(l)}.$$

Alors dans ce cas, on peut remplacer \mathcal{A} par \mathcal{B} : c'est-à-dire le problème du test entre les mesures μ_1, \dots, μ_n sur \mathcal{A} et équivalent à celui des mesures μ_1, \dots, μ_n sur \mathcal{B} .

Un résumé exhaustif nous est fourni par le théorème suivant (voir [F₂]).

THÉORÈME 1. — *Soit \mathcal{X} un espace de Banach, séparable et réflexif, et X un l -élément aléatoire à valeurs dans \mathcal{X} . Supposons que nous voulions tester par une ou plusieurs observations indépendantes de X , à laquelle des n lois μ_1, \dots, μ_n définies sur le σ -corps engendré par les ensembles « cylindriques » définis par les fonctionnelles linéaires fortement continues sur \mathcal{X} , obéit X . Si nous désignons par μ une mesure par rapport à laquelle μ_1, \dots, μ_n sont absolument continues et par $f_i(x)$ la densité de Radon-Nikodym de μ_i par rapport à μ , alors un résumé exhaustif pour le test en question nous est fourni par les $(n-1)$ rapports*

$$\frac{f_i(x)}{f_l(x)} \quad \text{pour tout } i \neq l.$$

l étant l'un des entiers $1, 2, \dots, n$.

En ce qui concerne l'absolue continuité des mesures μ_i les unes par rapport aux autres, on a le résultat suivant (voir [G₁] ou [F₂]) :

THÉORÈME 2. — *Désignons par s_j ($j = 1, 2, \dots, n$) les valeurs propres (distinctes ou non) et par $u_j(t)$ les fonctions propres, uni-*

taires, deux à deux orthogonales de l'opérateur linéaire dans \mathfrak{X} , symétrique et défini positif,

$$(1.4) \quad b(t) = \int_0^T \Gamma(t, \tau) a(\tau) d\tau$$

ou symboliquement $b = \Gamma(a)$.

Une condition nécessaire et suffisante pour que μ_α soit absolument continue par rapport à μ_0 est que $\rho_\alpha(t) \in \mathbf{H}$, ensemble des fonctions définies sur $[0, T]$ et de la forme

$$h(t) = \sum_{j=1}^{\infty} h_j u_j(t) \quad (0 \leq t \leq T).$$

avec

$$\sum_{j=1}^{\infty} \frac{(h_j)^2}{s_j} < +\infty.$$

De plus si $\rho_\alpha \in \mathbf{H}$, μ_0 est aussi absolument continue par rapport à μ_α et

$$\frac{d\mu_\alpha}{d\mu_0}(x) = e^{\psi_\alpha(x)},$$

où

$$\psi_\alpha(x) = \sum_{j=1}^{\infty} \frac{\rho_\alpha^j}{s_j} \left(x_j - \frac{\rho_\alpha^j}{2} \right),$$

avec

$$x^j = \int_0^T x(t) u_j(t) dt \quad \text{et} \quad \rho_\alpha^j = \int_0^T \rho_\alpha(t) u_j(t) dt.$$

2. — Cas stationnaire. Canal de communication de bande passante finie.

— En technique des communications, $U(t)$ représente le bruit de fond et les fonctions $\rho_i(t)$ des signaux. L'étude d'un canal bruyant ayant une bande passante fixée, finie (voir [S₁]) nous amène à considérer des covariances de la forme

$$(2.1) \quad \Gamma(t, \tau) = r(t - \tau) = \int_{-\Omega}^{+\Omega} e^{i\omega(t-\tau)} f(\omega) d\omega.$$

C'est-à-dire $U(t)$ est la partie relative à l'intervalle $[0, T]$ d'une fonction aléatoire $\tilde{U}(t)$ stationnaire sur $(-\infty, +\infty)$, de fonction de corrélation $r(h)$, de spectre nul en dehors d'une bande $(-\Omega, +\Omega)$ et admettant dans cette bande une densité spectrale $f(\omega)$.

De même, nous considérons des fonctions $\rho_\alpha(t)$ qui sont les parties relatives à $[0, T]$ de signaux réels définis sur $(-\infty, +\infty)$:

$$(2.2) \quad \tilde{\rho}_\alpha(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} \varphi_\alpha(\omega) d\omega.$$

Nous supposons que $f(\omega)$ est de carré sommable et différent de zéro sauf peut-être sur un ensemble de mesure de Lebesgue nulle sur $(-\Omega, +\Omega)$.

LEMME 1. — *Sous les hypothèses faites précédemment, les $u_j(t)$ forment une base complète dans \mathfrak{X} .*

En écrivant que

$$s_j u_j = \Gamma(u_j).$$

on obtient

$$s_j u_j(t) = \int_0^T \left[\int_{-\Omega}^{+\Omega} e^{i\omega t - i\omega\tau} f(\omega) d\omega \right] u_j(\tau) d\tau \quad (0 \leq t \leq T),$$

en intervertissant les signes d'intégration

$$(2.3) \quad = \int_{-\Omega}^{+\Omega} e^{i\omega t} g_j(\omega) f(\omega) d\omega,$$

où

$$(2.4) \quad g_j(\omega) = \int_0^T e^{-i\omega\tau} u_j(\tau) d\tau.$$

Si les u_j ne formaient pas une base complète dans \mathfrak{X} , on pourrait trouver une fonction $u(t)$, élément non nul de \mathfrak{X} , telle que

$$\int_0^T r(t-\tau) u(\tau) d\tau = 0 \quad \text{pour } 0 \leq t \leq T.$$

Mais

$$\begin{aligned} \int_0^T r(t-\tau) u(\tau) d\tau &= \int_0^T \left[\int_{-\Omega}^{+\Omega} e^{i\omega t - i\omega\tau} f(\omega) d\omega \right] u(\tau) d\tau \\ &= \int_{-\Omega}^{+\Omega} e^{i\omega t} f(\omega) \left[\int_0^T e^{-i\omega\tau} u(\tau) d\tau \right] d\omega. \end{aligned}$$

Donc, en posant

$$g(\omega) = \int_0^T e^{-i\omega\tau} u(\tau) d\tau,$$

on aurait

$$h(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} g(\omega) f(\omega) d\omega \equiv 0 \quad \text{pour } 0 \leq t \leq T.$$

La fonction $h(t)$ est une fonction entière de t ; étant nulle sur $[0, T]$ elle serait identiquement nulle. Par suite de l'unicité de la transformée de Fourier, $g(\omega) f(\omega)$ serait presque partout nulle sur $[-\Omega, +\Omega]$. Comme sur $(-\Omega, +\Omega)$, $f(\omega)$ est différent de zéro presque partout, $g(\omega)$ serait nulle presque partout sur $(-\Omega, +\Omega)$ et étant une fonction entière de ω , serait identiquement nulle, même pour $|\omega| > \Omega$; donc sa transformée de Fourier $u(t)$ est nécessairement l'élément nul de \mathfrak{X} .

Posons

$$(2.5) \quad u'_j(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} g_j(\omega) \sqrt{f(\omega)} d\omega \quad \text{pour } -\infty < t < +\infty$$

LEMME 2. — *Sur $(-\infty, +\infty)$, les $u'_j(t)$ sont orthogonales.*

En effet, calculons

$$\begin{aligned} (2.6) \quad & \int_{-\infty}^{+\infty} u'_j(t) \overline{u'_k(t)} dt \\ &= 2\pi \int_{-\Omega}^{+\Omega} g_j(\omega) \sqrt{f(\omega)} \overline{g_k(\omega)} \sqrt{f(\omega)} d\omega \\ &= 2\pi \int_{-\Omega}^{+\Omega} \left[\int_0^T \int_0^T e^{-i\omega\tau} u_j(\tau) e^{+i\omega\tau'} u_k(\tau') d\tau d\tau' \right] f(\omega) d\omega \\ &= 2\pi \int_0^T \int_0^T u_j(\tau) u_k(\tau') \left[\int_{-\Omega}^{+\Omega} e^{-i\omega\tau} e^{+i\omega\tau'} f(\omega) d\omega \right] d\tau d\tau' \\ &= 2\pi \int_0^T \int_0^T r(\tau - \tau') u_j(\tau) u_k(\tau') d\tau d\tau' \\ &= 2\pi s_j \int_0^T u_j(\tau') u_k(\tau') d\tau' = \begin{cases} 0 & \text{si } j \neq k. \\ 2\pi s_j & \text{si } j = k. \end{cases} \end{aligned}$$

En général, bien entendu, les $u'_j(t)$ ne forment pas une base complète dans l'espace \mathfrak{X}' des fonctions de carré sommable sur $(-\infty, +\infty)$.

Toutefois, on a

LEMME 3. — *Toute fonction de \mathfrak{X}' de la forme*

$$h(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} g(\omega) d\omega$$

avec $g(\omega)$ de carré sommable sur $(-\Omega, +\Omega)$, appartient au sous-espace de \mathfrak{X} engendré par les u_j .

En effet :

Il est équivalent de montrer que $(g_j(\omega)\sqrt{f(\omega)})$ constituent une base complète dans l'espace des fonctions de carré sommable sur $(-\Omega, +\Omega)$. Supposons qu'il n'en soit pas ainsi : il existerait $\theta(\omega)$ élément non nul de l'espace des fonctions de carré sommable sur $(-\Omega, +\Omega)$ telle que

$$\int_{-\Omega}^{+\Omega} \theta(\omega) \bar{g}_j(\omega) \sqrt{f(\omega)} d\omega = 0 \quad \text{pour tout } j.$$

Mais

$$\begin{aligned} \int_{-\Omega}^{+\Omega} \theta(\omega) \bar{g}_j(\omega) \sqrt{f(\omega)} d\omega &= \int_{-\Omega}^{+\Omega} \theta(\omega) \sqrt{f(\omega)} \left\{ \int_0^T e^{i\omega\tau} u_j(\tau) d\tau \right\} d\omega \\ &= \int_0^T u_j(\tau) \int_{-\Omega}^{+\Omega} e^{i\tau\omega} \theta(\omega) \sqrt{f(\omega)} d\omega. \end{aligned}$$

Comme les u_j forment une base pour \mathfrak{X} , alors

$$\int_{-\Omega}^{+\Omega} e^{i\tau\omega} \theta(\omega) \sqrt{f(\omega)} d\omega = 0 \quad \text{pour } 0 \leq t \leq T.$$

Mais étant entière, cette fonction serait nulle partout.

Par suite de l'unicité de la transformée de Fourier, on aurait

$$\theta(\omega) \sqrt{f(\omega)} = 0 \quad \text{presque partout pour } |\omega| \leq \Omega,$$

donc

$$\theta(\omega) \text{ lui-même} = 0 \quad \text{presque partout pour } |\omega| \leq \Omega.$$

3. Discrimination entre deux signaux. — Supposons que nous ayons à tester entre deux signaux seulement, nous pouvons alors supposer que l'un d'eux est identiquement nul, c'est-à-dire nous envisagerons le problème du test entre pas de signal et le signal $\rho(t)$ qui, rappelons-le, est la partie relative à $[0, T]$ de

$$(3.1) \quad \tilde{\rho}(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} \varphi(\omega) d\omega.$$

Supposons que

$$(3.2) \quad \int_{-\Omega}^{+\Omega} \frac{|\varphi(\omega)|^2}{f(\omega)} d\omega < +\infty.$$

Et posons

$$(3.3) \quad \rho'(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} \frac{\varphi(\omega)}{\sqrt{f(\omega)}} d\omega.$$

D'après le lemme 3,

$$\rho'(t) = \sum_{j=1}^{\infty} \lambda_j u_j'(t)$$

avec la topologie forte dans \mathfrak{X}' , de sorte que

$$\int_{-\infty}^{+\infty} |\rho'(t)|^2 dt = 2\pi \sum s_j |\lambda_j|^2 < +\infty.$$

Or on a

$$2\pi s_j \lambda_j = \int_{-\infty}^{+\infty} \rho'(t) \bar{u}_j'(t) dt = 2\pi \int_{-\Omega}^{+\Omega} \frac{\varphi(\omega)}{\sqrt{f(\omega)}} \bar{g}_j(\omega) \sqrt{f(\omega)} d\omega$$

d'après (2.5) et (3.3);

$$\begin{aligned} 2\pi s_j \lambda_j &= 2\pi \int_{-\Omega}^{+\Omega} \varphi(\omega) \left[\int_0^T e^{i\omega t} u_j(t) dt \right] d\omega \\ &= 2\pi \int_0^T \left[\int_{-\Omega}^{+\Omega} e^{i\omega t} \varphi(\omega) d\omega \right] u_j(t) dt \\ &= 2\pi \int_0^T \rho(t) u_j(t) dt = 2\pi \rho_j, \end{aligned}$$

d'où

$$(3.4) \quad \lambda_j = \frac{\rho_j}{s_j}$$

et par suite

$$(3.5) \quad 2\pi \sum_j \frac{(\rho_j)^2}{s_j} = \int_{-\infty}^{+\infty} |\rho'(t)|^2 dt = 2\pi \int_{-\Omega}^{+\Omega} \frac{|\varphi(\omega)|^2}{f(\omega)} d\omega.$$

La formule (3.5) nous donne une interprétation physique de la quantité $\sum_j \frac{(\rho_j)^2}{s_j}$ en la reliant directement à l'énergie totale du signal $\rho'(t)$ défini par (3.3).

D'autre part, on en déduit que $\rho(t)$ appartient à l'espace H défini au théorème 2 (§ 1).

On a de plus le théorème suivant :

THÉORÈME 3. — *Pour une covariance ayant la forme indiquée au paragraphe 2, toute fonction $\rho(t)$ qui est la partie relative à $[0, T]$ d'une fonction de la forme*

$$\int_{-\Omega}^{+\Omega} e^{i\omega t} \varphi(\omega) d\omega, \quad \text{avec} \quad \int_{-\Omega}^{+\Omega} \frac{|\varphi(\omega)|^2}{f(\omega)} d\omega < +\infty$$

appartient à H ; et réciproquement tout $h \in H$ est égal presque partout à une fonction de cette forme.

Nous venons de voir que de telles fonctions appartiennent à H . Réciproquement soit $h \in H$,

$$h(t) = \sum_j h_j u_j(t), \quad \text{avec} \quad \sum_j \frac{(h_j)^2}{s_j} < +\infty.$$

Posons

$$\eta(\omega) = \sum_j \frac{h_j}{s_j} g_j(\omega) \sqrt{f(\omega)},$$

c'est une fonction de carré sommable sur $(-\Omega, +\Omega)$.

Soit alors

$$\tilde{h}(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} \eta(\omega) \sqrt{f(\omega)} d\omega$$

et calculons

$$\begin{aligned} \int_0^T \tilde{h}(t) u_j(t) dt &= \int_0^T u_j(t) \left[\int_{-\Omega}^{+\Omega} e^{i\omega t} \eta(\omega) \sqrt{f(\omega)} d\omega \right] dt \\ &= \int_{-\Omega}^{+\Omega} \eta(\omega) \sqrt{f(\omega)} \left[\int_0^T e^{i\omega t} u_j(t) dt \right] d\omega \\ &= \int_{-\Omega}^{+\Omega} \eta(\omega) \sqrt{f(\omega)} \overline{g_j(\omega)} d\omega = h_j, \end{aligned}$$

donc $\tilde{h}(t) = h(t)$ presque partout sur $[0, T]$.

Revenons au test entre pas de signal et le signal $\rho(t)$ vérifiant (3.1) et (3.2). Les mesures μ_0 et μ , respectivement définies (comme nous l'avons vu au paragraphe 1), par les lois temporelles des fonctions aléatoires $U(t)$ et $U(t) + \rho(t)$ sont équivalentes.

Comme nous l'avons vu également au paragraphe 1, un résumé exhaustif par le test considéré, est constitué par l'expression

$$(3.6) \quad \log \frac{d\mu}{d\mu_0}(X) = \sum_{j=1}^{\infty} \frac{\rho^j}{s_j} \left(X^j - \frac{\rho^j}{2} \right),$$

avec

$$X^j = \int_0^T X(t) u_j(t) dt.$$

En utilisant ce résumé exhaustif pour la discrimination, le test devient particulièrement simple puisque $\sum_{j=1}^{\infty} \frac{\rho^j}{s_j} \left(X^j - \frac{\rho^j}{2} \right)$ obéit à une loi de Laplace de variance $A = \sum_j \frac{(\rho^j)^2}{s_j}$ et de moyenne $-\frac{A}{2}$ si X obéit à μ_0 et $+\frac{A}{2}$ si X obéit à μ .

Remarque. — Nous déduisons en particulier de ce qui précède que l'erreur de discrimination dépend uniquement de la quantité A , c'est-à-dire de l'énergie totale de $\rho'(t)$. Elle ne dépend donc pas de T . Ce qui pourrait sembler un paradoxe, résulte de propriétés d'analyticité que nous allons examiner.

Les fonctions

$$\tilde{\varphi}_\alpha(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} \varphi_\alpha(\omega) d\omega$$

sont analytiques sur toute la droite. De plus, presque sûrement, les trajectoires de la fonction aléatoire $\tilde{U}(t)$, stationnaire, de moyenne nulle et de fonction de corrélation $r(h) = \int_{-\Omega}^{+\Omega} e^{i\omega h} f(\omega) d\omega$ sont des fonctions analytiques.

Soit alors $T^* > T$ et posons

$$U^*(t) = \tilde{U}(t) \quad \text{et} \quad \rho_\alpha^*(t) = \tilde{\varphi}_\alpha(t) \quad \text{pour} \quad 0 \leq t \leq T^*.$$

Considérons le test de la valeur de α à partir de l'observation de

$$X^*(t) = \rho_\alpha^*(t) + U^*(t) \quad \text{pour} \quad 0 \leq t \leq T^*;$$

Nous avons $U(t) = U^*(t)$, $\rho_\alpha(t) = \rho_\alpha^*(t)$, $X(t) = X^*(t)$, pour $0 \leq t \leq T$; donc l'information apportée sur la valeur de α par l'observation de $X^*(t)$,

$0 \leq t \leq T^*$, est plus grande ou égale à celle fournie par l'observation de $X(t)$, $0 \leq t \leq T$. Mais par suite de l'analyticité des réalisations de $X^*(t)$ sur $[0, T^*]$, à chaque réalisation de $X(t)$ sur $[0, T]$ correspond une seule réalisation de $X^*(t)$ sur $[0, T^*]$; donc les deux informations précédentes sont égales, ce qui explique évidemment le fait que l'erreur de discrimination ne dépend pas de T , durée d'observation.

Cette constatation nous incline à penser que les hypothèses faites ne sont pas rigoureusement compatibles avec un cas concret.

Sans doute, dans la pratique, l'hypothèse d'un spectre limité strictement à une bande finie bien que couramment faite par les ingénieurs, n'est pas rigoureusement remplie, ne serait-ce qu'à cause de l'erreur introduite par l'inertie de tout système enregistreur, erreur qui peut toujours être rentrée dans la fonction aléatoire $U(t)$.

Néanmoins, il semble que les considérations théoriques développées ici peuvent fournir des indications utiles pour des cas rencontrés dans des problèmes pratiques. Il resterait évidemment à étudier les conditions de stabilité des résultats lorsque les hypothèses ne sont pas rigoureusement remplies [par exemple, si le spectre n'est pas entièrement nul sur $(-\Omega, +\Omega)$].

4. Discrimination entre plus de deux signaux. — Nous considérons dans ce paragraphe, le même « bruit » $U(t)$ que dans les paragraphes 2 et 3, et cette fois $(n+1)$ signaux $\rho_1(t) \dots \rho_{n+1}(t)$, sur chacun desquels nous supposons :

a Que $\rho_\alpha(t)$ est la partie relative à $[0, T]$ d'un signal réel $\tilde{\rho}_\alpha(t)$ sur $(-\infty, +\infty)$ de spectre nul en dehors de la bande $(-\Omega, +\Omega)$.

Soit

$$(4.1) \quad \tilde{\rho}_\alpha(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} \varphi_\alpha(\omega) d\omega.$$

b Qu'on a

$$(4.2) \quad \int_{-\Omega}^{+\Omega} \frac{|\varphi_\alpha(\omega)|^2}{f(\omega)} d\omega \leq R^2.$$

Remarque. — Dans le cas d'un bruit blanc, c'est-à-dire

$$f(\omega) = \begin{cases} = \text{Cte} & \text{pour } |\omega| \leq \Omega, \\ = 0 & \text{pour } |\omega| > \Omega. \end{cases}$$

On a

$$N = E \{ |U(t)|^2 \} = \int_{-\Omega}^{+\Omega} f(\omega) d\omega = 2\Omega \times \text{Cte},$$

d'où $\text{Cte} = \frac{N}{2\Omega}$ et

$$\begin{aligned} \int_{-\Omega}^{+\Omega} \frac{|\varphi_\alpha(\omega)|^2}{f(\omega)} d\omega &= \frac{2\Omega}{N} \int_{-\Omega}^{+\Omega} |\varphi_\alpha(\omega)|^2 d\omega \\ &= \frac{2\Omega}{2\pi N} \int_{-\infty}^{+\infty} |\tilde{\rho}_\alpha(t)|^2 dt. \end{aligned}$$

Donc, dans ce cas, l'hypothèse (4.2) est équivalente à la suivante : les signaux $\tilde{\rho}_\alpha(t)$ ont une énergie totale limitée par $\pi \frac{R^2 N}{\Omega}$; si nous désignons par E cette énergie, alors R^2 peut s'écrire

$$R^2 = \frac{\Omega E}{\pi N}.$$

Comme nous l'avons vu au paragraphe précédent, les conditions a et b imposées aux $\rho_\alpha(t)$, impliquent que $\rho_\alpha(t) \in H$. Donc les mesures μ_1, \dots, μ_{n+1} définies par la loi temporelle de $U(t) + \rho_1(t), \dots, U(t) + \rho_{n+1}(t)$ sont équivalentes; et alors pour le problème de la discrimination entre $\rho_1(t), \dots, \rho_{n+1}(t)$, à partir de l'observation de la fonction $X(t) = U(t) + \rho_k(t)$, (k étant inconnu), pour $0 \leq t \leq T$, nous pouvons utiliser le résumé exhaustif constitué par le système des n variables aléatoires $\psi_2, \dots, \psi_{n+1}$ suivantes :

$$(4.3) \quad \psi_\alpha = \sum_{j=1}^{\infty} \frac{(\rho_\alpha^j - \rho_1^j)}{s_j} \left[X_j - \frac{\rho_\alpha^j - \rho_1^j}{2} \right],$$

avec

$$(4.4) \quad \begin{cases} X_j = \int_0^T X(t) u_j(t) dt, \\ \rho_\alpha^j = \int_0^T \rho_\alpha(t) u_j(t) dt. \end{cases}$$

Or ces n variables aléatoires obéissent à une loi de Laplace de dimension n ; appelons \mathcal{L}_k cette loi dans le cas où c'est $\rho_k(t)$ qui est le signal d'émission.

En posant

$$(4.5) \quad B_{kl} = \sum_j \frac{(\rho_l^j - \rho_1^j)(\rho_k^j - \rho_1^j)}{s_j} = B_{lk}.$$

Le point moyen de \mathcal{L}_k se détermine par

$$E_k[\psi_\alpha] = B_{\alpha k} - \frac{1}{2} B_{\alpha\alpha} \quad (\alpha = 2, 3, \dots, n+1);$$

Quant à la matrice des covariances correspondantes à \mathcal{L}_k , elle a pour éléments les quantités $B_{\alpha\beta}$.

Donc toute modification \mathcal{M} du système $\rho_1(t), \dots, \rho_{n+1}(t)$ qui laisserait invariant le système des $B_{k,l}$, ne modifie rien aux conditions du test.

Notons que H est un espace de Hilbert avec le produit scalaire

$$[h', h''] = \sum_j \frac{h'^j h''^j}{s_j}.$$

Alors remarquons que, les conditions du test restent inchangées si $\rho_1, \dots, \rho_{n+1}$ sont remplacés par $n+1$ signaux $\rho_1^*, \dots, \rho_{n+1}^*$ déduits des ρ_α par un opérateur unitaire dans H et que la condition (4.2), qui s'écrit

$$(4.6) \quad \sum_j \frac{(\rho_\alpha^j)^2}{s_j} = \int_{-\Omega}^{+\Omega} \frac{|\varphi_\alpha(\omega)|^2}{f(\omega)} d\omega \leq R^2$$

demeure aussi inchangée par un tel opérateur.

Soit H_n le sous-espace de H constitué par les $h(t)$ telles que

$$h^j = 0 \quad \text{pour tout } j > n;$$

il existe toujours un opérateur unitaire \mathcal{G} dans H qui transforme $\rho_1, \dots, \rho_{n+1}$ en des éléments $\rho_1^*, \dots, \rho_{n+1}^*$ de H_n .

Donc, en ce qui concerne l'étude du test, on ne réduit pas la généralité en supposant que

$$\rho_\alpha \in H_n \quad \text{pour tout } \alpha = 1, 2, \dots, n+1.$$

Dans ces conditions, les ψ_α ne dépendent que de X^1, \dots, X^n et par suite on peut prendre comme résumé exhaustif ces variables aléatoires ou mieux :

$$\left(\frac{X^1}{\sqrt{s_1}}, \dots, \frac{X^n}{\sqrt{s_n}} \right),$$

système de n variables aléatoires, laplaciennes, indépendantes, de variance 1 et dont le point moyen, si c'est $\rho_k(t)$ qui a été émis, vérifie

$$E\left(\frac{X^j}{\sqrt{s_j}}\right) = \frac{\rho_k^j}{\sqrt{s_j}};$$

et (4.6) s'écrit

$$\sum_{j=1}^n \frac{(\rho_k^j)^2}{s_j} \leq R^2.$$

Nous sommes ramenés alors au problème suivant :

Étant donné un espace Euclidien de dimension n , E_n , rapporté à des axes de coordonnées rectangulaires, soient $n + 1$ points M_1, \dots, M_{n+1} dans la sphère de centre O et de rayon R , et Y^j des variables aléatoires laplaciennes, indépendantes de variance 1.

Désignons par $\{m_k^j\}$ les coordonnées du point M_k . Alors par une observation sur le point L de coordonnées

$$m_k^j + Y^j \quad \text{où } k \text{ est inconnu mais fixé,}$$

nous voulons déterminer la valeur de k .

Pour cela, il nous faut partager l'espace E_n en $(n + 1)$ régions disjointes S_1, S_2, \dots, S_{n+1} telles que si L tombe dans S_i , on décide que $k = i$. Si, par exemple, les différentes valeurs de k sont *a priori* équiprobables et si ν_i désigne la loi de probabilité dans E_n de L quand $k = i$, l'erreur totale est

$$\frac{1}{n + 1} \sum_{i=1}^{n+1} \nu_i(\bar{S}_i) \quad (\bar{S}_i = \text{complémentaire de } S_i \text{ dans } E_n).$$

On sait (voir [A₁]) que si A_{ij}, \bar{A}_{ij} est une décomposition de Hahn de E_n , par rapport à la mesure algébrique $(\nu_i - \nu_j)$, alors le système de régions qui minimise cette erreur est donné par

$$S_i = \bigcap_{j \neq i} A_{ij}.$$

Par suite, dans le cas présent, on est ramené à considérer les plans médiateurs des segments $M_i M_j$. La région d'acceptation de M_i sera la portion d'espace délimitée par les plans médiateurs des segments $M_i M_j$ pour tout j qui contient le point M_i .

Soit alors \mathcal{E} l'erreur totale de discrimination correspondante aux meilleures régions que nous venons de définir; \mathcal{E} dépend évidemment de R , de n , et de la disposition \mathcal{O} des points M_k dans la sphère de centre O et de rayon R .

Soit

$$\mathcal{E}(R, n, \mathcal{O});$$

R et n étant fixés, on peut considérer le minimum de \mathcal{E} pour toutes les dispositions \mathcal{O} de points M_k dans la sphère de centre O et de rayon R .

Soit

$$\mathcal{E}_0(R, n) = \min_{\mathcal{O}} \mathcal{E}(R, n, \mathcal{O});$$

il est alors intéressant d'introduire, pour $\varepsilon > 0$ fixé, $n(R, \varepsilon)$ le plus grand entier n tel que

$$\mathcal{E}_0(R, n) \leq \varepsilon.$$

§. Capacité d'un canal de communication perturbée par un bruit blanc laplacien. — Considérons une émission de signaux réels, dont le spectre est limité à la bande $(-\Omega, +\Omega)$, c'est-à-dire de la forme

$$(§.1) \quad \tilde{\varphi}(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} \varphi(\omega) d\omega.$$

Si P désigne la puissance moyenne du dispositif émetteur, alors une durée de fonctionnement de T , permettra d'obtenir tout signal $\tilde{\varphi}(t)$ d'énergie totale inférieure à PT :

$$(§.2) \quad \int_{-\infty}^{+\infty} |\tilde{\varphi}(t)|^2 dt \leq PT.$$

Nous supposons qu'un récepteur reçoive ce signal perturbé par un bruit Laplacien $U(t)$ dont le spectre est limité à la même bande $(-\Omega, +\Omega)$; c'est-à-dire qu'à l'instant t , il reçoit :

$$(§.3) \quad X(t) = \tilde{\varphi}(t) + U(t),$$

et nous prendrons pour période d'observation l'intervalle $0 \leq t \leq T$ par exemple.

On désire alors tester quel est le signal qui a été émis, sachant que celui-ci appartient à un certain ensemble de signaux possibles. C'est le problème étudié précédemment.

Nous supposons de plus que $U(t)$ est un bruit blanc, c'est-à-dire, comme nous l'avons vu,

$$f(\omega) = \begin{cases} \frac{N}{2\Omega} & \text{pour } |\omega| \leq \Omega, \\ 0 & \text{pour } |\omega| > \Omega, \end{cases}$$

où $N = E(|U(t)|^2)$ désigne la puissance moyenne du bruit.

D'autre part, soient $\tilde{\rho}_1, \dots, \tilde{\rho}_{n+1}$ les différents signaux possibles

$$\tilde{\rho}_\alpha(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} \varphi_\alpha(\omega) d\omega$$

et $\rho_1, \dots, \rho_{n+1}$ leurs parties relatives à $[0, T]$.

La condition (5.2) s'écrit

$$(5.4) \quad \int_{-\Omega}^{+\Omega} \frac{|\varphi_\alpha(\omega)|^2}{f(\omega)} d\omega = \frac{\Omega}{\pi N} \int_{-\infty}^{+\infty} |\tilde{\rho}_\alpha(t)|^2 dt \leq \frac{\Omega PT}{\pi N}$$

Supposons que les différents signaux possibles soient *a priori* équiprobables. Quel est alors le nombre maximal de signaux parmi lesquels on peut discriminer avec une erreur totale inférieure à ε ?

C'est précisément

$$n\left(\sqrt{\frac{\Omega PT}{\pi N}}, \varepsilon\right).$$

L'étude du maximum de la vitesse de transmission de l'information à travers un tel dispositif avec une erreur aussi petite que possible, nous amène à considérer

$$(5.5) \quad \lim_{\varepsilon \rightarrow 0} \lim_{T \rightarrow \infty} \frac{\log n\left(\sqrt{\frac{\Omega PT}{\pi N}}, \varepsilon\right)}{T}.$$

Remarque. — Notons que la double limite précédente, si elle existe, sera nécessairement une fonction de $\Omega \frac{P}{N}$.

Par suite, la formule avancée par Shannon [S₁] pour la capacité d'un canal de communication ayant la bande passante $(-\Omega, +\Omega)$, la puissance moyenne de l'émetteur étant P et celle du bruit N :

$$C = \Omega \log\left(1 + \frac{P}{N}\right)$$

est incompatible avec les considérations précédentes.

Intuitivement, il semble que $\mathcal{E}_0(R, n)$ soit atteint pour des points M_1, \dots, M_{n+1} placés aux sommets d'un polyèdre régulier inscrit dans la sphère de centre O et de rayon R, tracée dans l'espace euclidien de dimension n .

Toutefois, une démonstration rigoureuse de ce fait est plus délicate qu'on ne pourrait le penser au premier abord et nous ne sommes pas parvenus à en établir une.

De toutes manières, soit

$$\mathcal{E}_1(\mathbf{R}, n)$$

l'erreur de discrimination correspondante à des points \mathbf{M}_k ainsi placés dans l'espace euclidien de dimension n aux sommets d'un polyèdre régulier inscrit dans la sphère de centre \mathbf{O} et de rayon \mathbf{R} .

On a évidemment

$$\mathcal{E}_1(\mathbf{R}, n) \geq \mathcal{E}_0(\mathbf{R}, n),$$

ε étant un nombre positif arbitraire, soit $n_1(\mathbf{R}, \varepsilon)$ le plus grand entier n tel que

$$\mathcal{E}_1(\mathbf{R}, n) \leq \varepsilon.$$

Nécessairement,

$$n_1(\mathbf{R}, \varepsilon) \leq n(\mathbf{R}, \varepsilon).$$

Nous allons étudier

$$\lim_{\varepsilon \rightarrow 0} \lim_{T \rightarrow \infty} \frac{\log n_1\left(\sqrt{\frac{\Omega PT}{\pi N}}, \varepsilon\right)}{T}.$$

Tout d'abord, nous allons effectuer trois calculs préliminaires qui nous seront utiles dans la suite :

1° Soit Z , une variable aléatoire laplacienne, réduite et a , un nombre positif

$$\begin{aligned} \Pr(Z \geq a) &= \frac{1}{\sqrt{2\pi}} \int_a^{+\infty} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{(x+a)^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{a^2}{2}} \int_0^{+\infty} e^{-\frac{x^2}{2}} e^{-xa} dx \\ &< \frac{1}{\sqrt{2\pi}} e^{-\frac{a^2}{2}} \int_0^{+\infty} e^{-xa} dx = \frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{a^2}{2}}}{a}. \end{aligned}$$

2° Soit toujours Z une variable aléatoire laplacienne réduite et a , un nombre positif

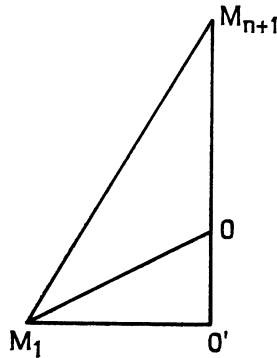
$$\Pr(Z \geq a) = \frac{1}{\sqrt{2\pi}} \int_a^{+\infty} e^{-\frac{x^2}{2}} dx.$$

Remarquons que $(x+1)e^{-x} < 1$ pour $x > 0$.

D'où

$$\Pr(Z \geq a) > \frac{1}{\sqrt{2\pi}} \int_a^{+\infty} e^{-\left(\frac{x^2}{2} + x\right)} (x + 1) dx = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{a^2}{2} + a\right)}.$$

3° Dans l'espace euclidien de dimension n soient M_1, \dots, M_{n+1} les sommets d'un polyèdre régulier inscrit dans la sphère de centre O et de rayon R . Calculons la distance δ_{n+1} entre deux sommets M_i et M_j .



Soit O' le centre de gravité de M_1, \dots, M_n . Les points O', O, M_{n+1} sont sur une même droite et le triangle $M_1O'M_{n+1}$ est rectangle.

On a

$$\overrightarrow{nO'O} = \overrightarrow{OM_{n+1}}$$

et

$$\begin{aligned} R^2 &= (OO')^2 + (O'M_1)^2, \\ (M_1M_{n+1})^2 &= (O'M_{n+1})^2 + (O'M_1)^2, \\ (M_1M_{n+1})^2 &= R^2 + R^2 \left(\frac{n+1}{n}\right)^2 - \frac{R^2}{n^2} = R^2 \left(1 + \frac{(n+1)^2 - 1}{n^2}\right) = 2R^2 \frac{n+1}{n}. \end{aligned}$$

Donc

$$\delta_{n+1}^2 = (M_iM_j)^2 = 2R^2 \frac{n+1}{n}.$$

Prenons donc les points M_1, \dots, M_{n+1} aux sommets d'un polyèdre régulier inscrit dans la sphère de centre O et de rayon R , tracée dans l'espace euclidien de dimension n .

Conformément à ce que nous avons vu plus haut, nous déciderons que le point moyen est M_i si L tombe dans la portion d'espace S_i délimitée par les plans médiateurs des segments M_iM_h pour tout h , qui contient le point M_i .

Par raison de symétrie, on a

$$v_1(S_1) = \dots = v_{n+1}(S_{n+1});$$

et par suite

$$\mathcal{E}_1 = \frac{1}{n+1} \sum_{i=1}^{n+1} v_i(\bar{S}_i) = v_i(\bar{S}_i).$$

Comme les points M_j sont sur une même sphère de centre O , dire que L appartient à S_i est équivalent aux n inégalités suivantes

$$\overrightarrow{OL} \cdot \overrightarrow{OM}_i > \overrightarrow{OL} \cdot \overrightarrow{OM}_h \quad \text{pour tout } h \neq i;$$

alors \mathcal{E}_1 est égale à la probabilité pour que, sachant que M_i est le point moyen, il existe au moins un $h \neq i$ tel que

$$\overrightarrow{OL} \cdot \overrightarrow{OM}_h > \overrightarrow{OL} \cdot \overrightarrow{OM}_i.$$

Soit

$$Z_h = \overrightarrow{OL} \cdot \overrightarrow{OM}_h,$$

Z_h est une variable aléatoire laplacienne;

$$\begin{aligned} E(Z_h) &= (OM_i)^2 = R^2, & \text{si } h = i, \\ E(Z_h) &= \overrightarrow{OM}_h \cdot \overrightarrow{OM}_i = -\frac{R^2}{n}, & \text{si } h \neq i, \end{aligned}$$

En effet :

$$\sum_{h=1}^{n+1} \overrightarrow{OM}_h = 0 \quad \text{et} \quad \overrightarrow{OM}_h \cdot \overrightarrow{OM}_i = \overrightarrow{OM}_{h'} \cdot \overrightarrow{OM}_i \quad \text{pour tous } h \text{ et } h' \neq i;$$

Donc

$$0 = \overrightarrow{OM}_i \cdot \sum_{h=1}^{n+1} \overrightarrow{OM}_h = R^2 + n \overrightarrow{OM}_h \cdot \overrightarrow{OM}_i.$$

Quant à la variance de Z_h , elle vaut

$$E\left(\overrightarrow{M_i L} \cdot \overrightarrow{OM}_h\right)^2 = E\left(\sum_{j=1}^n Y_j m_k^j\right)^2 = \sum_{j=1}^n (m_k^j)^2 = R^2.$$

Au moins pour R assez grand, on peut trouver λ compris entre 0 et R tel que

$$\Pr(Z_i - R^2 < -\lambda R) < \frac{\varepsilon}{2};$$

D'autre part, soit $h \neq i$:

$$\begin{aligned} \Pr[Z_h > R(R - \lambda)] &= \Pr\left(\frac{Z_h + \frac{R^2}{n}}{R} > (R - \lambda) + \frac{R}{n}\right) \\ &\leq \Pr\left(\frac{Z_h + \frac{R^2}{n}}{R} > (R - \lambda)\right) \leq \frac{1}{\sqrt{2\pi}(R - \lambda)} e^{-\frac{(R - \lambda)^2}{2}}. \end{aligned}$$

Alors $\Pr(Z_i > Z_h \text{ pour tout } h \neq i) \geq$

$$\begin{aligned} &\Pr[Z_i > R(R - \lambda) \text{ et tous les } Z_h \text{ avec } h \neq i < R(R - \lambda)] \\ &= 1 - \Pr(Z_i < R(R - \lambda) \text{ ou au moins l'un des } Z_h > R(R - \lambda) \text{ pour } h \neq i) \\ &\geq 1 - \Pr(Z_i < R(R - \lambda)) - n \Pr(Z_h > R(R - \lambda)) \\ &\geq 1 - \frac{\varepsilon}{2} - \frac{n}{\sqrt{2\pi}(R - \lambda)} e^{-\frac{(R - \lambda)^2}{2}}. \end{aligned}$$

Donc

$$\begin{aligned} \mathcal{E}_1(R, n) &= \Pr \text{ il existe au moins un } h \neq i \text{ tel que } Z_h > Z_i \\ &\leq \frac{\varepsilon}{2} + \frac{n}{\sqrt{2\pi}(R - \lambda)} e^{-\frac{(R - \lambda)^2}{2}}. \end{aligned}$$

Par suite, si

$$\frac{n}{\sqrt{2\pi}(R - \lambda)} e^{-\frac{(R - \lambda)^2}{2}} \leq \frac{\varepsilon}{2},$$

c'est-à-dire

$$n \leq \frac{\varepsilon}{2} \sqrt{2\pi}(R - \lambda) e^{\frac{(R - \lambda)^2}{2}}.$$

On a

$$\mathcal{E}_1(R, n) \leq \varepsilon,$$

donc,

$$n \leq n_1(R, \varepsilon),$$

$$\begin{aligned} \frac{1}{T} \log n_1\left(\sqrt{\frac{\Omega PT}{\pi N}}, \varepsilon\right) &\geq \frac{1}{T} \log \frac{\varepsilon}{2} \sqrt{2\pi} \\ &\quad + \frac{1}{T} \log\left(\sqrt{\frac{\Omega PT}{\pi N}} - \lambda(\varepsilon)\right) + \frac{1}{2T} \left(\sqrt{\frac{\Omega PT}{\pi N}} - \lambda(\varepsilon)\right)^2 \end{aligned}$$

et, par suite,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \log n_1\left(\sqrt{\frac{\Omega PT}{\pi N}}, \varepsilon\right) \geq \frac{\Omega P}{2\pi N};$$

donc

$$(3.6) \quad \lim_{\varepsilon \rightarrow 0} \lim_{T \rightarrow \infty} \frac{1}{T} \log n_1\left(\sqrt{\frac{\Omega PT}{\pi N}}, \varepsilon\right) \geq \frac{\Omega P}{2\pi N}.$$

THÉOREME 4. — *Quels que soient ε et η positifs aussi petits qu'on veut, pour T assez grand, on peut transmettre à travers le canal considéré pendant le temps T une quantité d'informations égale à $\left(\frac{\Omega P}{2\pi N} - \eta\right) T$ avec une fréquence d'erreur inférieure à ε .*

Nous allons maintenant obtenir une majoration pour

$$\overline{\lim}_{T \rightarrow \infty} \frac{1}{T} \log n_1 \left(\sqrt{\frac{\Omega P}{\pi N}} T, \varepsilon \right).$$

Cette majoration serait surtout intéressante si l'on avait montré que $n(R, \varepsilon) = n_1(R, \varepsilon)$ auquel cas elle fournirait une borne supérieure pour la vitesse de transmission de l'information dans le canal.

Supposons que le point moyen soit M_1 et posons

$$W_{i-1} = Z_i - Z_1 = \overrightarrow{OL} \cdot \overrightarrow{M_1 M_i}.$$

Nous décidons que M_1 est le point moyen si $W_i < 0$ pour $i = 1, \dots, n$; W_1, \dots, W_n obéissent conjointement à une loi de Laplace à n dimensions :

$$\begin{aligned} E(W_i) &= E(Z_i) - E(Z_1) = -\frac{R^2}{n} - R^2 = -\frac{n+1}{n} R^2 = -\frac{\delta_{n+1}^2}{2}, \\ E(W_i - EW_i)(W_k - EW_k) &= E(\overrightarrow{M_1 L} \cdot \overrightarrow{M_1 M_i})(\overrightarrow{M_1 L} \cdot \overrightarrow{M_1 M_k}) \\ &= E\left(\sum_{j=1}^n (m_j^i - m_j^1) Y_j\right) \left(\sum_{j=1}^n (m_j^k - m_j^1) Y_j\right) = \sum_{j=1}^n (m_j^i - m_j^1)(m_j^k - m_j^1) \\ &= (\overrightarrow{OM_i} - \overrightarrow{OM_1}) \cdot (\overrightarrow{OM_k} - \overrightarrow{OM_1}) = \begin{cases} \frac{\delta_{n+1}^2}{2} & \text{si } i = k, \\ \frac{\delta_{n+1}^2}{2} & \text{si } i \neq k. \end{cases} \end{aligned}$$

Considérons la loi de W_n quand W_1, \dots, W_{n-1} sont fixés.

Remarquons que :

$$\begin{aligned} \xi_1 &= W_1, \\ \xi_2 &= W_1 - 2W_2, \\ &\dots\dots\dots \\ \xi_n &= W_1 + \dots + W_{n-1} - nW_n \end{aligned}$$

sont des variables aléatoires mutuellement indépendantes.

Calculons

$$\begin{aligned} E(\xi_n) &= (n-1) \left(-\frac{\delta^2}{2}\right) - n \left(-\frac{\delta^2}{2}\right) = \frac{\delta^2}{2}; \\ \text{Variance}(\xi_n) &= E[(W_1 - EW_1) + (W_2 - EW_2) + \dots \\ &\quad + (W_{n-1} - EW_{n-1}) - n(W_n - EW_n)]^2 \\ &= (n-1)\delta^2 + n^2\delta^2 + 2\delta_{n-1}^2 \frac{\delta^2}{2} - 2n(n-1) \frac{\delta^2}{2} \\ &= \delta^2 \left(n-1 + n^2 + \frac{(n-1)!}{(n-3)!2} - n(n-1) \right) \\ &= \delta^2 \left(2-1 + \frac{(n-1)(n-2)}{2} \right) = \frac{\delta^2}{2} n(n+1). \end{aligned}$$

Fixer W_1, \dots, W_{n-1} revient à fixer ξ_1, \dots, ξ_{n-1} . Donc la loi de ξ_n lorsque W_1, \dots, W_{n-1} sont fixés est la même que sa loi marginale et par suite, comme

$$W_n = \frac{W_1 + \dots + W_{n-1}}{n} - \frac{\xi_n}{n},$$

la loi de W_n quand W_1, \dots, W_{n-1} sont fixés est une loi de Laplace de moyenne

$$\begin{aligned} \frac{w_1 + \dots + w_{n-1}}{n} - \frac{\delta^2}{2n} \quad \text{et de variance} \quad \frac{\delta^2(n-1)}{2n}; \\ \mathcal{E}_1(R, n) &= \Pr \quad (\text{au moins un des } W_i > 0) \\ &\geq \Pr \quad (\text{un des } W_i \text{ est positif et les autres négatifs}) \\ &= n \Pr \quad (W_n > 0, W_1 < 0, \dots, W_{n-1} < 0). \end{aligned}$$

Posons

$$\begin{aligned} V_i &= \frac{W_i - EW_i}{\sigma(W_i)} = \frac{W_i + \frac{\delta^2}{2}}{\delta}, \\ W_i &= \delta V_i - \frac{\delta^2}{2}. \end{aligned}$$

Soit $f(v_1, \dots, v_{n-1})$ la densité de la loi conjointe de V_1, \dots, V_{n-1} .

On a

$$\begin{aligned} \Pr(W_n > 0, W_1 < 0, \dots, W_{n-1} < 0) \\ &= \int_{-\infty}^{\frac{\delta}{2}} \dots \int_{-\infty}^{\frac{\delta}{2}} \Pr(W_n > 0 / v_1, \dots, v_{n-1}) f(v_1, \dots, v_{n-1}) dv_1, \dots, dv_{n-1} \\ &= \int_{-\infty}^{\frac{\delta}{2}} \dots \int_{-\infty}^{\frac{\delta}{2}} \Pr\left(\frac{W_n - m_c}{\sigma_c} > -\frac{m_c}{\sigma_c} \mid v_1, \dots, v_{n-1}\right) f(v_1, \dots, v_{n-1}) dv_1, \dots, dv_{n-1}, \end{aligned}$$

où m_c et σ_c désignent respectivement la moyenne conditionnelle et l'écart type conditionnel de W_n quand V_1, \dots, V_{n-1} sont fixés. Mais :

$$m_c = \frac{w_1 + \dots + w_{n-1}}{n} - \frac{\delta^2}{2n} = \delta \frac{v_1 + \dots + v_{n-1}}{n} - \frac{\delta^2}{2},$$

$$\sigma_c = \delta \sqrt{\frac{n+1}{2n}}.$$

en employant une inégalité établie précédemment :

$$\Pr\left(\frac{W_n - m_c}{\sigma_c} > -\frac{m_c}{\sigma_c} \mid v_1, \dots, v_{n-1}\right)$$

$$\geq \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{n}{n+1} \left(\frac{v_1 + \dots + v_{n-1}}{n} - \frac{\delta}{2}\right)^2\right]$$

$$+ \sqrt{\frac{2n}{n+1}} \left(\frac{v_1 + \dots + v_{n-1}}{n} - \frac{\delta}{2}\right),$$

d'où

$$\Pr(W_n > 0, W_1 < 0, \dots, W_{n-1} < 0)$$

$$\geq \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{n}{n+1} \frac{\delta^2}{4} - \sqrt{\frac{2n}{n+1}} \frac{\delta}{2}\right]$$

$$\times \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} \dots \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} \exp\left[-\frac{(v_1 + \dots + v_{n-1})^2}{n(n+1)}\right]$$

$$+ \frac{n}{n+1} \delta \frac{v_1 + \dots + v_{n-1}}{n}$$

$$+ \sqrt{\frac{2n}{n+1}} \left(\frac{v_1 + \dots + v_{n-1}}{n}\right)]$$

$$\times f(v_1, \dots, v_{n-1}) dv_1, \dots, dv_{n-1}$$

$$\geq \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{n}{n+1} \frac{\delta^2}{4} - \sqrt{\frac{2n}{n+1}} \frac{\delta}{2} - \frac{(n-1)^2}{n(n+1)^2} \frac{\delta^2}{4}\right]$$

$$- \frac{n}{n+1} \frac{n-1}{n} \frac{\delta^2}{2} - \sqrt{\frac{2n}{n+1}} \frac{n-1}{n} \frac{\delta}{2}]$$

$$\int_{-\frac{\delta}{2}}^{+\frac{\delta}{2}} \dots \int_{-\frac{\delta}{2}}^{+\frac{\delta}{2}} f(v_1, \dots, v_{n-1}) dv_1, \dots, dv_{n-1}$$

$$\geq \frac{1}{\sqrt{2\pi}} \exp(-\delta^2 - \delta\sqrt{2}) \int_{-\frac{\delta}{2}}^{+\frac{\delta}{2}} \dots \int_{-\frac{\delta}{2}}^{+\frac{\delta}{2}} f(v_1, \dots, v_{n-1}) dv_1, \dots, dv_{n-1}.$$

Mais

$$\begin{aligned} & \int_{-\frac{\delta}{2}}^{+\frac{\delta}{2}} \cdots \int_{-\frac{\delta}{2}}^{+\frac{\delta}{2}} f(v_1, \dots, v_{n-1}) dv_1, \dots, dv_{n-1} \\ &= \Pr \left(\text{tous les } V_i \text{ soient compris entre } -\frac{\delta}{2} \text{ et } +\frac{\delta}{2} \right) \\ &= 1 - \Pr \left(\text{au moins un des } V_i > \frac{\delta}{2} \text{ ou } < -\frac{\delta}{2} \right) \\ &\geq 1 - 2 \Pr \left(\text{au moins un des } V_i > \frac{\rho}{2} \right) \\ &= 1 - 2 \Pr \left(\text{au moins un des } W_i > 0, i = 1, \dots, n-1 \right) \\ &\geq 1 - 2 \Pr \left(\text{au moins un des } W_i > 0, i = 1, \dots, n \right); \end{aligned}$$

donc

$$\int_{-\frac{\delta}{2}}^{+\frac{\delta}{2}} \cdots \int_{-\frac{\delta}{2}}^{+\frac{\delta}{2}} f(v_1, \dots, v_{n-1}) dv_1, \dots, dv_{n-1} \geq 1 - 2 \mathcal{E}_1(R, n).$$

Et par suite,

$$\mathcal{E}_1(R, n) \geq \frac{n}{\sqrt{2\pi}} [1 - 2 \mathcal{E}_1(R, n)] \cdot \exp(-\delta^2 - \delta \sqrt{2}).$$

Si nous prenons $n = n_1(R, \varepsilon)$, nous avons

$$\varepsilon \geq \mathcal{E}_1(R, n_1(R, \varepsilon)) \geq \frac{n_1(R, \varepsilon)}{\sqrt{2\pi}} [1 - 2\varepsilon] \exp(-\delta^2 - \delta \sqrt{2})$$

donc

$$n_1(R, \varepsilon) \leq \frac{\sqrt{2\pi} \varepsilon}{1 - 2\varepsilon} \exp(\delta^2 + \delta \sqrt{2}).$$

Mais

$$\begin{aligned} \delta &= \delta_{n+1} = \sqrt{2} R \sqrt{\frac{n+1}{n}}, \\ n_1(R, \varepsilon) &\leq \frac{\sqrt{2\pi} \varepsilon}{1 - 2\varepsilon} \exp\left(2 \frac{n+1}{n_1} R^2 + 2 \sqrt{\frac{n+1}{n_1}} R\right). \end{aligned}$$

On sait que $n_1(R, \varepsilon)$ augmente indéfiniment avec R . Donc, quel que soit $\eta > 0$, pour R suffisamment grand on aura

$$n_1(R, \varepsilon) \leq \frac{\sqrt{2\pi} \varepsilon}{1 - 2\varepsilon} \exp(2(I + \eta) R^2 + 2 \sqrt{I + \eta} R).$$

Et l'on a

$$\lim_{T \rightarrow \infty} \frac{1}{T} \log n_1\left(\sqrt{\frac{\Omega PT}{\pi N}}, \varepsilon\right) \leq (I + \eta) \frac{2 \Omega P}{\pi N}$$

quel que soit $\eta > 0$; donc,

$$\overline{\lim}_{T \rightarrow \infty} \frac{1}{T} \log n_1 \left(\sqrt{\frac{\Omega PT}{\pi N}}, \varepsilon \right) \leq \frac{2\Omega P}{\pi N},$$

et par suite,

$$\lim_{\varepsilon \rightarrow 0} \overline{\lim}_{T \rightarrow \infty} \frac{1}{T} \log n_1 \left(\sqrt{\frac{\Omega PT}{\pi N}}, \varepsilon \right) \leq \frac{2\Omega P}{\pi N}.$$

CHAPITRE II.

ESTIMATION PAR LE MAXIMUM DE VRAISEMBLANCE DE LA MOYENNE D'UNE FONCTION ALÉATOIRE PERMANENTE.

1. Position du problème et notations. — Ayant effectué certaines observations sur une fonction aléatoire $X(t)$ [en l'occurrence on supposera qu'on a enregistré toutes les valeurs de $X(t)$ pour t parcourant un certain intervalle], nous envisagerons dans ce chapitre le problème de l'estimation de la valeur moyenne $E[X(t)]$ sachant que cette fonction appartient à un certain ensemble de fonctions données.

Plus précisément supposons que nous ayons observé pour $0 \leq t \leq T$ une fonction aléatoire de la forme

$$(1.1) \quad X(t) = U(t) + \rho(t) \quad (0 \leq t \leq T),$$

où $U(t)$ est une fonction aléatoire, réelle, telle que $E[U(t)] = 0$ et $\rho(t)$ est une fonction réelle, certaine, mais inconnue, dont on sait seulement qu'elle appartient à une famille donnée \mathcal{F} de fonctions définies sur $[0, T]$.

Un estimateur de $\rho(t)$ sera un procédé qui à toute réalisation de $X(t)$ (c'est-à-dire ensemble des valeurs de $X(t)$ pour tout $t \in [0, T]$) associe une fonction de \mathcal{F} présumée être la moyenne de $X(t)$ ou tout au moins assez près (en un sens à préciser) de cette moyenne. Faisons sur la fonction aléatoire $U(t)$ les hypothèses suivantes :

- 1° $U(t)$ est laplacienne;
- 2° sa covariance $\Gamma(t, \tau)$ est continue et connue.

Supposons de plus que \mathcal{F} soit contenu dans \mathcal{X} espace de Hilbert des fonctions réelles de la variable t sur l'intervalle $[0, T]$ qui sont de carré sommable sur cet intervalle par rapport à la mesure de Lebesgue

Comme nous l'avons vu dans le chapitre précédent,

$$(1.2) \quad V_\rho(t) = U(t) + \rho(t)$$

peut-être considéré comme un élément aléatoire V_ρ à valeurs dans \mathfrak{X} . Soit μ_ρ la mesure correspondante à la loi temporelle de $V_\rho(t)$ définie sur le plus petit σ -corps \mathfrak{A} de sous-ensembles de \mathfrak{X} qui rend mesurable les applications

$$x \xrightarrow{\omega_{t(\cdot)}} x(t)$$

de $(\mathfrak{X}, \mathfrak{A})$ dans (\mathbb{R}, ζ) pour tout $t \in [0, T]$.

μ_0 désignera la mesure correspondante à $U(t)$.

$s_j, u_j(t), H$ auront la même signification que dans le théorème 2 du paragraphe 1 du chapitre I;

Nous noterons par

$$(1.3) \quad \langle g_1, g_2 \rangle = \int_0^T g_1(t) g_2(t) dt$$

le produit scalaire habituel dans \mathfrak{X} .

Alors une condition nécessaire et suffisante pour que μ_ρ soit absolument continue par rapport à μ_0 est que ρ appartienne à H . Et dans ce cas, on a

$$\frac{d\mu_\rho}{d\mu_0}(x) = \exp \sum_{j=1}^{\infty} \frac{\rho^j}{s_j} \left(x^j - \frac{\rho^j}{s_j} \right)$$

presque sûrement, avec

$$\begin{aligned} \rho^j &= \langle \rho, u_j \rangle, \\ x^j &= \langle x, u_j \rangle. \end{aligned}$$

Nous supposerons dans la suite que \mathfrak{F} est contenu dans H qui, remarquons-le est un espace de Hilbert si on le munit du produit scalaire

$$(1.4) \quad [h_1, h_2] = \sum_{j=1}^{\infty} \frac{h_1^j h_2^j}{s_j},$$

où $h_i^j = \langle h_i, u_j \rangle$.

L'exemple des problèmes classiques d'estimations dans le cas de nombres aléatoires, nous porte à essayer ici, également, une méthode du maximum de vraisemblance.

2. Estimateur du maximum de vraisemblance. — Pour ρ fixé, $\frac{d\mu_\rho}{d\mu_0}$ densité de μ_ρ par rapport à μ_0 définit en réalité une classe de fonctions (deux

fonctions de la classe étant égales presque partout). Supposons donc que, d'une certaine manière, nous ayons fait choix, pour chaque $\rho \in \mathcal{F}$, d'une fonction $f_\rho(x)$ dans la classe $\frac{d\mu_\rho}{d\mu_0}$; alors s'il existe une application $\hat{\rho}$ de \mathcal{X} dans \mathcal{F} telle que pour tout $x \in \mathcal{X}$

$$(2.1) \quad f_{\hat{\rho}}(x)^{(x)} \geq f_\rho(x) \quad \text{pour tout } \rho \in \mathcal{F},$$

Nous appellerons $\hat{\rho}$, *estimateur du maximum de vraisemblance*.

Pour faire le choix, évoqué plus haut, nous pouvons songer à prendre

$$f_\rho(x) = \exp \sum_{j=1}^{\infty} \frac{\rho^j}{s_j} \left(x^j - \frac{\rho^j}{2} \right)$$

partout où la série $\sum_{j=1}^{\infty} \frac{\rho^j}{s_j} \left(x^j - \frac{\rho^j}{2} \right)$ converge et $f_\rho(x) = \text{Cte}$ par exemple 1, là où elle ne converge pas.

Nous savons que pour tout $\rho \in \mathbb{H}$, $\sum_{j=1}^{\infty} \frac{\rho^j x^j}{s_j}$ converge presque sûrement par rapport à n'importe quelle mesure $\mu_{\rho'}$, $\rho' \in \mathbb{H}$. (En effet, on a affaire à une somme de variables aléatoires, indépendantes, chacune de moyenne nulle, dont la somme des variances converge). C'est-à-dire il existe $E_\rho \in \mathcal{A}$ tel que

$$\begin{aligned} 1^\circ \quad & \mu_{\rho'}(E_\rho) = 0 \quad \text{pour tout } \rho' \in \mathbb{H}. \\ 2^\circ \quad & \sum_j \frac{\rho^j x^j}{s_j} \text{ converge pour tout } x \notin E_\rho. \end{aligned}$$

Mais on ne peut affirmer sans autres hypothèses que presque sûrement,

$$\sum_{j=1}^{\infty} \frac{\rho^j x^j}{s_j} \text{ converge pour tout } \rho \in \mathcal{F};$$

c'est-à-dire que l'ensemble exceptionnel peut être pris indépendant de ρ .

Supposons que \mathcal{V} soit un sous-espace vectoriel fermé (par rapport à la topologie induite par le produit scalaire (1,4) de \mathbb{H} , alors nous pouvons énoncer le lemme suivant :

LEMME 1. — Soit \mathcal{V} un sous-espace vectoriel fermé de \mathbb{H} . Si pour $x \in \mathcal{X}$.

$$\sum_{j=1}^{\infty} \frac{\rho^j x_j}{s_j} \text{ converge pour tout } \rho \in \mathcal{V},$$

alors $\sum_{j=1}^{\infty} \frac{\rho^j x^j}{s_j}$ est une fonctionnelle linéaire fortement continue sur \mathfrak{V} muni de la topologie induite par le produit scalaire (1.4).

En effet,

$$\mathfrak{T}(\rho) = \sum_{j=1}^{\infty} \frac{\rho^j x^j}{s_j} = \lim_{n \rightarrow \infty} \sum_{j=1}^n \frac{\rho^j x^j}{s_j}.$$

Pour n fixé, soit

$$\begin{aligned} \mathfrak{T}_n(\rho) &= \sum_{j=1}^n \frac{\rho^j x^j}{s_j}, \\ |\mathfrak{T}_n(\rho)| &\leq \sum_{j=1}^n \frac{|\rho^j| \cdot |x^j|}{s_j}. \end{aligned}$$

Mais

$$\|\rho\|_{\mathfrak{H}}^2 = \sum_{j=1}^{\infty} \frac{(\rho^j)^2}{s_j},$$

où

$$\frac{(\rho^j)^2}{s_j} \leq \|\rho\|_{\mathfrak{H}}^2.$$

Et

$$|\mathfrak{T}_n(\rho)| \leq \sum_{j=1}^n \frac{|x^j|}{\sqrt{s_j}} \|\rho\|_{\mathfrak{H}} = \left(\sum_{j=1}^n \frac{|x^j|}{\sqrt{s_j}} \right) \|\rho\|_{\mathfrak{H}}.$$

Donc, $\mathfrak{T}_n(\rho)$ est une fonctionnelle linéaire, fortement continue sur \mathfrak{H} par suite sur $\mathfrak{V} \subset \mathfrak{H}$.

Et alors $\mathfrak{T}(\rho) = \lim_{n \rightarrow \infty} \mathfrak{T}_n(\rho)$ est aussi une fonctionnelle linéaire fortement continue sur \mathfrak{V} (voir [D₂] théorème 17, p. 54).

Conséquence. — Supposons remplies les conditions du lemme 1 : Alors il existe $\hat{\rho}_x \in \mathfrak{V}$ tel que

$$[\rho, \hat{\rho}_x] = \sum_{j=1}^{\infty} \frac{\rho^j x^j}{s_j} \text{ pour tout } \rho \in \mathfrak{V}.$$

Et pour tout $\rho \in \mathfrak{V}$,

$$(2.2) \quad \sum_{j=1}^{\infty} \frac{\rho^j}{s_j} \left(x^j - \frac{\hat{\rho}_x^j}{2} \right) = [\rho, \hat{\rho}_x] - \frac{1}{2} \|\rho\|_{\mathfrak{H}}^2 \leq \frac{1}{2} \|\hat{\rho}_x\|_{\mathfrak{H}}^2 = \sum_{j=1}^{\infty} \frac{\hat{\rho}_x^j}{s_j} \left(x^j - \frac{\hat{\rho}_x^j}{2} \right).$$

En effet l'inégalité n'est autre que $\|\rho - \hat{\rho}_x\| \geq 0$.

Considérant toujours le cas où $\mathcal{F} = \mathcal{V}$ sous-espace vectoriel fermé de H , la propriété précédente montre l'importance de l'ensemble \mathcal{E} des x tel que

$$\sum_{j=1}^{\infty} \frac{\rho^j x^j}{s_j} \text{ converge pour tout } \rho \in \mathcal{V}.$$

Au sujet de cet ensemble, on peut montrer le théorème suivant, dont je dois la majeure partie à A. Hanen.

THÉORÈME 1. — *Si \mathcal{V} est de dimension finie, alors*

$$\mu_0(\mathcal{E}) = 1 = \mu_{\rho}(\mathcal{E});$$

si \mathcal{V} est de dimension infinie, alors,

$$\mu_0(\mathcal{E}) = 0 = \mu_{\rho}(\mathcal{E}).$$

1° Supposons que \mathcal{V} soit de dimension infinie et soit $\{\xi_k\}$ une base orthonormale, complète de \mathcal{V} .

$\rho \in \mathcal{V}$ s'écrit

$$\rho = \sum_{k=1}^{\infty} [\rho, \xi_k] \xi_k, \quad \text{avec } \sum_{k=1}^{\infty} [\rho, \xi_k]^2 < +\infty.$$

Pour $x \in \mathcal{E}$ et $\rho \in \mathcal{V}$, posons

$$Z(x, \rho) = \sum_{j=1}^{\infty} \frac{\rho^j x^j}{s_j},$$

$$\rho = \sum_{k=1}^K [\rho, \xi_k] \xi_k + \varepsilon(K), \quad \|\varepsilon(K)\|_H \rightarrow 0 \quad \text{quand } K \rightarrow \infty,$$

$$Z(x, \rho) = \sum_{j=1}^{\infty} \frac{x^j}{j} \sum_{k=1}^K [\rho, \xi_k] \xi_k^j + Z(x, \varepsilon(K)),$$

$$\begin{aligned} \sum_{j=1}^{\infty} \frac{x^j}{s_j} \sum_{k=1}^K [\rho, \xi_k] \xi_k^j &= \lim_{J \rightarrow \infty} \sum_{j=1}^J \frac{x^j}{s_j} \sum_{k=1}^K [\rho, \xi_k] \xi_k^j \\ &= \lim_{J \rightarrow \infty} \sum_{k=1}^K [\rho, \xi_k] \sum_{j=1}^J \frac{x^j \xi_k^j}{s_j} = \sum_{k=1}^K [\rho, \xi_k] Z(x, \xi_k). \end{aligned}$$

Comme $x \in \mathcal{E}$,

$$Z(x, \rho) = \sum_{k=1}^K [\rho, \xi_k] Z(x, \xi_k) + Z(x, \varepsilon(K)).$$

En faisant tendre K vers l'infini et par suite de la continuité forte en ρ de $Z(x, \rho)$ on a

$$Z(x, \rho) = \sum_{k=1}^{\infty} [\rho, \xi_k] Z(x, \xi_k)$$

Mais alors, dire que pour $x \in \mathcal{E}$, $Z(x, \rho)$ est une fonctionnelle linéaire continue sur \mathcal{V} entraîne que :

$$(2.3) \quad \sum_{k=1}^{\infty} |Z(x, \xi_k)|^2 < +\infty.$$

Donc, pour $x \in \mathcal{E}$,

$$(2.4) \quad \sum_{k=1}^{\infty} \left| \sum_{j=1}^{\infty} \frac{x^j \xi_k^j}{s_j} \right|^2 < +\infty.$$

Or $\sum_{j=1}^{\infty} \frac{x^j \xi_k^j}{s_j}$ sont des variables aléatoires, laplaciennes, indépendantes de moyenne nulle (si x obéit à la loi μ_0) et de variance 1; donc, d'après la loi des grands nombres

$$\frac{\sum_{k=1}^n \left| \sum_{j=1}^{\infty} \frac{x^j \xi_k^j}{s_j} \right|^2}{n} \rightarrow 1 \quad \text{presque sûrement } \mu_0.$$

En comparant avec (2.4), on a

$$\mu_0(\mathcal{E}) = 0.$$

2° Supposons que \mathcal{V} soit de dimension finie,

$$\rho \in \mathcal{V} \quad \text{entraîne} \quad \rho = \sum_{k=1}^N \alpha_k \xi_k.$$

Soit E_k l'ensemble de x , où $\sum_{j=1}^{\infty} \frac{x^j \xi_k^j}{s_j}$ ne converge pas, si $E = \bigcup_{k=1}^N E_k$,

on a

$$\mu_0(E) \leq \sum_{k=1}^N \mu_0(E_k) = 0,$$

et si $x \in \bigcap E = \mathcal{E}$,

$$\sum_{j=1}^{\infty} \frac{\rho^j x^j}{s_j} = \sum_{k=1}^N \alpha_k \sum_{j=1}^{\infty} \frac{\xi_k^j x^j}{s_j}$$

converge pour tout $\rho \in \mathcal{V}$.

Ce qui démontre le théorème énoncé.

3. Cas où \mathcal{F} est un sous-espace vectoriel de dimension finie de H . — \mathcal{F} est l'ensemble des éléments de la forme

$$(3.1) \quad \rho(t) = \sum_{k=1}^N l_k \xi_k(t),$$

où les ξ_k sont N éléments linéairement indépendants de H , que nous ne supposons plus nécessairement orthogonaux.

Alors si \mathcal{E} est l'ensemble introduit dans le théorème 1, pour tout $\rho \in \mathcal{F}$, nous pouvons prendre d'après ce théorème, pour densité de μ_ρ par rapport à μ_0 :

$$(3.2) \quad f_\rho(x) = \exp \sum_{j=1}^{\infty} \frac{\rho^j}{s_j} \left(x^j - \frac{\rho^j}{2} \right) \quad \text{si } x \in \mathcal{E},$$

$$= 1 \quad \text{si } x \notin \mathcal{E}.$$

Pour $x \in \mathcal{E}$, soit

$$\hat{\rho}_x = \sum_{k=1}^N \hat{l}_k(x) \xi_k$$

tel que

$$(3.3) \quad [\rho, \hat{\rho}_x] = \sum_{j=1}^{\infty} \frac{\rho^j x^j}{s_j} \quad \text{pour tout } \rho \in \mathcal{F};$$

$\hat{\rho}_x$ existe d'après le lemme 1, et les \hat{l}_k sont donnés par le système de Cramer :

$$(3.4) \quad \sum_{j=1}^{\infty} \frac{\xi_l^j x^j}{s_j} = \sum_{k=1}^N \hat{l}_k [\xi_l, \xi_k] \quad (l = 1, \dots, N).$$

THÉORÈME 2. — $\hat{\rho}$ est un estimateur du maximum de vraisemblance.

Nous n'avons défini $\hat{\rho}$ que pour $x \in \mathcal{E}$; définissons-le de façon arbitraire pour $x \notin \mathcal{E}$.

En prenant $f_{\hat{\rho}}(x)$ comme indiqué par (3.2), on a

$$f_{\hat{\rho}}(x) \geq f_\rho(x) \quad \text{pour tout } \rho \in \mathcal{F}$$

[pour $x \in \mathcal{E}$, cette inégalité résulte de l'inégalité (2.2)].

THÉORÈME 3. — $\hat{\rho}$ est un estimateur exhaustif et non biaisé.

Le fait que $\hat{\rho}$ soit non biaisé, se constate immédiatement en prenant les espérances mathématiques dans le deux termes des équations (3.4)

et en se souvenant que par hypothèse le déterminant $\{[\xi_i, \xi_j]\}$ est différent de zéro.

En ce qui concerne l'exhaustivité, nous remarquerons qu'il existe une vraie probabilité (ou régulière) conditionnelle de \mathbf{X} lorsque $\hat{\rho}$ est fixé, puisque \mathcal{X} est un espace de Hilbert séparable et que $\overline{\mathcal{A}}$ complétion de \mathcal{A} est égale à $\overline{\mathcal{B}}$ complétion du σ -corps des ensembles de Borel de \mathcal{X} , comme nous l'avons vu au début du chapitre I, il suffit alors d'appliquer le corollaire III de [J₁].

Pour constater que la probabilité conditionnelle de \mathbf{X} lorsque $\hat{\rho}$ est fixé ne dépend pas de ρ , il suffit de remarquer que pour $x \in \mathcal{E}$:

$$\log f_{\rho}(x) = [\rho, \hat{\rho}_x]_{\mathbf{H}} - \frac{1}{2} \|\rho\|_{\mathbf{H}}^2.$$

Alors le résultat se déduit de l'application du théorème 1 de [H₂].

THÉORÈME 4. — *Pour chaque t fixé, $\hat{\rho}_x(t)$ valeur de $\hat{\rho}_x$ au point t , est parmi toutes les estimations non biaisées de $\rho(t)$ à variance minimale.*

Ce théorème est mentionné sans démonstration explicite dans [S₂]. Fixons t .

Soit $u(x)$ une estimation de $\rho(t)$ (valeurs de ρ au point t); nous dirons qu'elle est non biaisée si

$$E_{\rho} [u(x)] = \rho(t) \quad \text{pour tout } \rho \in \mathcal{F},$$

où E_{ρ} indique qu'on prend l'espérance mathématique par rapport à la loi μ_{ρ} .

Pour une estimation non biaisée de $\rho(t)$, on a alors

$$\text{var}_{\rho} u(x) = E_{\rho} [u(x)]^2 - [\rho(t)]^2,$$

et $u(x)$ sera à variance uniformément minimale, si, pour toute autre estimation non biaisée $u'(x)$ on a

$$E_{\rho} [u(x)]^2 \leq E_{\rho} [u'(x)]^2 \quad \text{pour tout } \rho \in \mathcal{F}.$$

Remarquons que, \mathbf{X} et \mathbf{Y} étant deux variables aléatoires quelconques, on a

$$E |Y|^2 \geq E |E(Y/X)|^2,$$

où $E(Y/X)$ indique l'espérance mathématique de \mathbf{Y} conditionnelle lors-

que X est fixé; en effet, il suffit de prendre l'espérance mathématique dans les deux termes de l'inégalité

$$E(|Y|^2/X) \leq |E(Y/X)|^2$$

qui résulte du fait qu'une variance est toujours non négative. Ceci étant, notons \hat{l} le vecteur $(\hat{l}_1, \dots, \hat{l}_N)$ et considérons

$$E_\rho[u(x)/\hat{l}]$$

C'est une fonction de \hat{l} indépendante de ρ puisque \hat{l} est un résumé exhaustif; c'est également une estimation non biaisée de $\rho(t)$, qui a une variance moindre que $u(x)$.

Donc, s'il existe une estimation non biaisée à variance uniformément minimale, elle appartiendra nécessairement à la classe des estimations qui sont des fonctions de \hat{l} , donc qu'on peut écrire

$$f(\hat{\rho}) = \hat{\rho}(t) + g(\hat{l}_1, \dots, \hat{l}_N).$$

Une telle estimation est non biaisée si et seulement si

$$E_\rho[g(\hat{l}_1, \dots, \hat{l}_N)] = 0 \quad \text{pour tout } l_1, \dots, l_N.$$

Nous supposons pour simplifier l'écriture que nous avons pris ξ_1, \dots, ξ_N orthogonales dans H , ce qui ne restreint pas la généralité. Alors $\hat{l}_1, \dots, \hat{l}_N$ sont des variables aléatoires laplaciennes, indépendantes.

On a

$$\int_{-\infty}^{+\infty} g(z_1, \dots, z_N) e^{-\frac{1}{2}\left(\frac{z_1^2}{\sigma_1^2} + \dots + \frac{z_N^2}{\sigma_N^2}\right)} e^{-\left(\frac{l_1 z_1}{\sigma_1^2} + \dots + \frac{l_N z_N}{\sigma_N^2}\right)} dz_1 \dots dz_N = 0$$

pour tout l_1, \dots, l_N ;

$$\int_{-\infty}^{+\infty} e^{-\frac{l_1 z_1}{\sigma_1^2}} e^{-\frac{1}{2}\frac{z_1^2}{\sigma_1^2}} \times \left[\int_{-\infty}^{+\infty} g(z_1, \dots, z_N) e^{-\frac{1}{2}\left(\frac{z_2^2}{\sigma_2^2} + \dots + \frac{z_N^2}{\sigma_N^2}\right)} e^{-\left(\frac{l_2 z_2}{\sigma_2^2} + \dots + \frac{l_N z_N}{\sigma_N^2}\right)} dz_2 \dots dz_N \right] dz_1 = 0$$

pour tout l_1 .

Par suite de l'unicité de la transformée de Laplace bilatérale, on a pour presque tout z_1

$$\int_{-\infty}^{+\infty} g(z_1, \dots, z_N) e^{-\frac{1}{2}\left[\frac{z_2^2}{\sigma_2^2} + \dots + \frac{z_N^2}{\sigma_N^2}\right]} e^{-\left(\frac{l_2 z_2}{\sigma_2^2} + \dots + \frac{l_N z_N}{\sigma_N^2}\right)} dz_2, \dots, dz_N = 0$$

pour tout l_2, \dots, l_N .

Et réitérant l'opération N fois, on obtient des ensembles A_1, \dots, A_N tous de mesure de Lebesgue nulle tels que si $z_i \notin A_i$ pour $i = 1, \dots, N$

$$g(z_1, \dots, z_N) = 0.$$

La loi de Laplace étant absolument continue par rapport à la mesure de Lebesgue, on a

$$g(\hat{z}_1, \dots, \hat{z}_N) = 0 \quad \text{presque sûrement.}$$

D'où

$$f(\hat{\rho}) = \hat{\rho}(t) \quad \text{presque sûrement } \mu_\rho, \forall \rho \in \mathcal{F}$$

et par suite $\hat{\rho}(t)$ est bien à variance uniformément minimale parmi toutes les estimations non biaisées de $\rho(t)$,

L'utilisation pratique des résultats précédents nécessite le calcul des quantités

$$[\xi_k, \xi_h] \quad \text{et} \quad \sum_{j=1}^{\infty} \frac{\rho^j x^j}{s_j}.$$

Dans plusieurs cas particuliers importants, on peut donner l'expression de ces quantités.

1. PROCESSUS DE WIENER-LÉVY, $\Gamma(t, \tau) = \min(t, \tau)$. — Supposons que les fonctions $\xi_k(t)$ s'annulent pour $t = 0$ et soient dérivables et à dérivée de carré sommable et à variation bornée sur l'intervalle $[0, T]$; alors

$$[\xi_k, \xi_h] = \int_0^T \xi'_k(t) \xi'_h(t) dt,$$

$$\sum_{j=1}^{\infty} \frac{\xi'_k x^j}{s_j} = \int_0^T \xi'_k(t) dx(t).$$

2. PROCESSUS DE ORNSTEIN-UHLENBECK, $\Gamma(t, \tau) = e^{-\beta|t-\tau|}$. — Supposons que les fonctions $\xi_k(t)$ soient dérivables et à dérivée de carré sommable et à variation bornée sur $[0, T]$;

$$\begin{aligned} [\xi_k, \xi_h] &= \frac{1}{2} [\xi_k(T) \xi_h(T) + \xi_k(0) \xi_h(0)] \\ &\quad + \frac{1}{2\beta} \int_0^T \xi'_k(t) \xi'_h(t) dt + \frac{\beta}{2} \int_0^T \xi_k(t) \xi_h(t) dt, \\ \sum_{j=1}^{\infty} \frac{\xi'_k x^j}{s_j} &= \frac{1}{2} [\xi_k(T) x(T) + \xi_k(0) x(0)] \\ &\quad + \frac{1}{2\beta} \int_0^T \xi'_k(t) dx(t) + \frac{\beta}{2} \int_0^T \xi_k(t) x(t) dt. \end{aligned}$$

3. PROCESSUS AUTO-RÉGRESSIF. — Plus généralement, en supposant que les fonctions $\xi_k(t)$ sont dérivables un nombre de fois suffisant, on peut donner des expressions (voir [P₁]) pour les quantités

$$[\xi_k, \xi_h] \text{ et } \sum_{j=1}^{\infty} \frac{\xi_k^j x^j}{s_j}$$

dans le cas où $U(t)$ est un processus dit auto-régressif (voir [D₁]).

4. SUPPOSONS QUE LES FONCTIONS $\xi_k(t)$ SOIENT DES ÉLÉMENTS DE $\Gamma(\mathcal{X})$; c'est-à-dire, il existe des fonctions τ_1, \dots, τ_N appartenant à \mathcal{X} telles que

$$\xi_k = \Gamma(\tau_k).$$

Alors

$$[\xi_k, \xi_h] = \langle \tau_k, \xi_h \rangle = \langle \xi_k, \tau_h \rangle, \quad \sum_{j=1}^{\infty} \frac{\xi_k^j x^j}{s_j} = \langle \tau_k, x \rangle.$$

On peut le voir facilement de la façon suivante :

Adjoignons au besoin aux u_j d'autres éléments $v_{j'}$ de \mathcal{X} , de façon à ce que l'ensemble constitue une base orthonormale complète de \mathcal{X} , alors tout élément $\eta \in \mathcal{X}$, s'écrit

$$\eta = \sum_i \tau_i^j u_j + \sum_{j'} \tau_i^{j'} v_{j'}.$$

En remarquant que les $v_{j'}$ sont pour l'opérateur Γ des fonctions propres correspondantes à la valeur propre 0, on a

$$\Gamma(\eta) = \sum_j s_j \tau_i^j u_j.$$

D'où

$$\langle \Gamma(\eta), u_j \rangle = s_j \langle \eta, u_j \rangle \quad \text{et} \quad \langle \Gamma(\eta), v_{j'} \rangle = 0.$$

Il s'ensuit les résultats énoncés plus haut.

Problèmes d'approximation. — Dans le cas d'un processus de Wiener-Lévy, et de celui d'un processus de Ornstein-Uhlenbeck, le système d'équations linéaires (3.4) donnant l'estimation \hat{p} a été obtenu par Mann (voir [M₁]) en considérant les équations du maximum de vraisemblance dans le cas où l'on observe $X(t)$ seulement en un nombre fini de points de l'intervalle $[0, T]$, puis en faisant tendre vers zéro la distance maximale entre deux consécutifs de ces points.

Plus généralement, on peut déduire des résultats de [S₂] que si l'on obtient des équations-limite par la méthode précédente, alors ces équations-limite ne sont autres que le système d'équations (3.4). Un autre genre de théorème d'approximation nous est fourni par Hanen (voir [H₃]). Le théorème 3 de [H₃] nous indique que :

1° si $\xi_i \in \Gamma(\mathcal{X})$ pour $i = 1, \dots, N$;

2° si $P_m(\mathcal{F})$ projection, au sens du produit scalaire $\langle . . . \rangle$ de \mathcal{X} , de \mathcal{F} sur \mathcal{X}_m sous-espace vectoriel de \mathcal{X} engendré par u_1, \dots, u_m , a même dimension que \mathcal{F} pour tout $m > N$, alors $\hat{\rho}_m$ estimation du maximum de vraisemblance à partir de $\langle X, u_i \rangle$ pour $i = 1, \dots, m$, converge au sens de la topologie de H (correspondant au produit scalaire $[\cdot; \cdot]$) vers $\hat{\rho}$ estimateur du maximum de vraisemblance à partir de X.

Cas non-laplacien. — On peut se demander ce que deviennent les estimations \hat{l}_k de l_k qui sont données par les équations (3.4) lorsqu'on abandonne le caractère laplacien du processus U(t).

On a le résultat suivant :

THÉOREME 5. — *Parmi toutes les estimations linéaires non-biaisées de l_k , à partir de X(t) pour $0 \leq t \leq T$, \hat{l}_k est à variance minimale.*

Considérons \mathcal{H} l'espace de Hilbert des variables aléatoires ayant un second moment fini, muni du produit scalaire

$$(v_1, v_2) = E(v_1 \bar{v}_2).$$

Soit alors E_k , l'ensemble des éléments de la forme

$$\sum_{i=1}^m d_i X(t_i) \quad \text{avec } t_i \in [0, T] \quad \text{et} \quad \sum_{i=1}^m d_i \xi_i(t_i) = \delta_{kl}.$$

Une estimation linéaire non biaisée de l_k , est un élément de la fermeture de E_k dans \mathcal{H} . C'est un ensemble convexe et fermé dans \mathcal{H} ; donc il existe une et une seule estimation linéaire non biaisée, à variance minimale l_k^* .

Les l_k^* sont caractérisés par l'équation matricielle

$$\begin{pmatrix} E[l_1^* X(t)] \\ \vdots \\ E[l_N^* X(t)] \end{pmatrix} = K \begin{pmatrix} \xi_1(t) \\ \vdots \\ \xi_N(t) \end{pmatrix} \quad \text{pour tout } 0 \leq t \leq T,$$

où K est une matrice constante. [voir $[C_1]$ en remarquant que l'hypothèse de stationnarité qui y est faite pour $U(t)$, n'est pas nécessaire. On peut remarquer également que les raisonnements sont semblables à ceux que nous ferons dans le paragraphe 3 du chapitre III.]

Alors on vérifie immédiatement que les \hat{l}_k vérifient une équation de ce type. D'où $l_k^* = \hat{l}_k$.

On a également le théorème suivant :

THÉORÈME 6. — Soit $t_0 \in [0, T]$ fixé, parmi toutes les estimations linéaires non biaisées de $\rho(t_0)$,

$$\hat{\rho}(t_0) = \sum_{k=1}^N \hat{l}_k \xi_k(t_0)$$

est à variance minimale.

Les estimations linéaires non biaisées de $\rho(t_0)$ sont les éléments de la fermeture dans \mathcal{H} de E ensemble des éléments de la forme :

$$\sum_{i=1}^m d_i X(t_i) \quad \text{avec} \quad t_i \in [0, T] \quad \text{et} \quad \sum_{i=1}^m l_i \xi_k(t_i) = \xi_k(t_0) \quad \text{pour} \quad k = 1, \dots, N.$$

\bar{E} est un ensemble fermé et convexe; donc il existe une et une seule estimation linéaire non biaisée de $\rho(t_0)$ qui soit à variance minimale : $\rho^*(t_0)$. Par des raisonnements tout à fait semblables à ceux que nous effectuerons pour le problème de prévision linéaire, étudié au chapitre III, on montre que $\rho^*(t_0)$ est caractérisé par la relation

$$E[\rho(t_0) - \rho^*(t_0)] X(t) = \sum_{k=1}^N C_k \xi_k(t) \quad \text{pour tout} \quad 0 \leq t \leq T,$$

où C_k sont des constantes.

$\hat{\rho}(t_0)$ vérifie une équation de ce type, d'où $\rho^*(t_0) = \hat{\rho}(t_0)$.

4. Quelques autres cas. — Nous avons vu dans le paragraphe précédent comment le cas où \mathcal{F} est un sous-espace vectoriel de dimension finie se prête à une étude complète et facile. Auparavant le théorème 1 nous avait montré la différence essentielle entre le cas d'un sous-espace vectoriel de dimension finie et celui d'un sous-espace vectoriel de dimension infinie et dans ce dernier cas les difficultés qui surgissent.

Dans le présent paragraphe nous étudierons trois autres cas où l'on peut définir un estimateur du maximum de vraisemblance.

1. \mathcal{F} SOUS-ENSEMBLE CONVEXE ET FERMÉ D'UN SOUS-ESPACE VECTORIEL DE DIMENSION FINIE DE H . — Soit \mathcal{F}_1 le sous-espace vectoriel de H engendré par \mathcal{F} .

Nous avons vu qu'il existe un ensemble $\mathcal{E} \in \mathcal{A}$ tel que $\mu_\rho(\mathcal{E}) = 1$ et une application $\hat{\rho}$ de \mathcal{E} dans \mathcal{F}_1 telle que

$$\sum_{j=1}^{\infty} \frac{\rho^j x^j}{s_j} = [\rho, \hat{\rho}_x] \quad \text{pour tout } \rho \in \mathcal{F}_1.$$

En définissant $f_\rho(x)$ par (3.2), nous avons pour $x \in \mathcal{E}$,

$$\log f_\rho(x) = \sum_{j=1}^{\infty} \frac{\rho^j}{s_j} \left(x^j - \frac{\rho^j}{2} \right) = [\rho, \hat{\rho}_x] - \frac{1}{2} [\rho, \rho] \quad \text{pour tout } \rho \in \mathcal{F}.$$

Mais alors rendre maximale cette expression est équivalent à rendre minimale

$$[\rho, \rho] + [\hat{\rho}_x, \hat{\rho}_x] - 2[\rho, \hat{\rho}_x] = [\rho - \hat{\rho}_x, \rho - \hat{\rho}_x] = \|\rho - \hat{\rho}_x\|_H;$$

\mathcal{F} étant supposé convexe et fermé dans H , il existe un et un seul élément $\bar{\rho}_x$ de \mathcal{F} tel que

$$\|\bar{\rho}_x - \hat{\rho}_x\|_H = \min_{\rho \in \mathcal{F}} \|\rho - \hat{\rho}_x\|_H;$$

et par suite

THÉORÈME 7. — *Il existe une et une seule application $\bar{\rho}$ de \mathcal{E} dans \mathcal{F} telle que*

$$f_{\bar{\rho}_x}(x) \geq f_\rho(x) \quad \text{pour tout } \rho \in \mathcal{F}.$$

Donc $\bar{\rho}_x$ est un estimateur du maximum de vraisemblance qu'on peut définir de n'importe quelle manière sur $\mathcal{X} - \mathcal{E}$.

2. \mathcal{F} EST COMPACT ET CONVEXE DANS $\Gamma(\mathcal{X})$. — Remarquons que $\rho \in \Gamma(\mathcal{X})$ entraîne $\rho = \sum_j \rho^j u_j$, avec $\sum_j \left(\frac{\rho^j}{s_j}\right)^2 < +\infty$; $\Gamma(\mathcal{X})$ peut être considéré comme un espace de Hilbert avec comme produit scalaire $\{\rho_1, \rho_2\} = \sum_j \frac{\rho_1^j \rho_2^j}{s_j^2}$.

Nous supposons que \mathcal{F} est un sous-ensemble de $\Gamma(\mathcal{X})$ compact par rapport à la topologie définie par $\{ \cdot, \cdot \}$. Si $\rho \in \Gamma(\mathcal{X})$, alors $\sum_j \frac{\rho^j x^j}{s_j}$ converge pour tout $x \in \mathcal{X}$ et

$$\sum_j \frac{\rho^j x^j}{s_j} \leq \|x\|_{\mathcal{X}} \left[\sum_j \left(\frac{\rho^j}{s_j} \right)^2 \right]^{\frac{1}{2}}.$$

Nous poserons

$$(4.1) \quad f_{\rho}(x) = \exp \sum_{j=1}^{\infty} \frac{\rho^j}{s_j} \left(x^j - \frac{\rho^j}{2} \right) \quad \text{pour tout } x \in \mathcal{X} \text{ et } \rho \in \mathcal{F};$$

on a

$$\psi(\rho, x) = \log f_{\rho}(x) \leq \|x\|_{\mathcal{X}} \left[\sum_j \left(\frac{\rho^j}{s_j} \right)^2 \right]^{\frac{1}{2}};$$

donc

$$M(x) = \sup_{\rho \in \mathcal{F}} \psi(\rho, x) < +\infty,$$

puisque nous avons supposé \mathcal{F} compact, donc borné dans $\Gamma(\mathcal{X})$.

Soit alors $\{\rho_n\}$ une suite d'éléments de \mathcal{F} telle que $\psi(\rho_n, x) \rightarrow M(x)$ quand $n \rightarrow \infty$; de $\{\rho_n\}$ on peut extraire une suite convergente vers un élément de \mathcal{F} , c'est-à-dire il existe $\{n_k\}$ et $\hat{\rho}_x \in \mathcal{F}$ tels que

$$\begin{aligned} \|\rho_{n_k} - \hat{\rho}_x\|_{\Gamma(\mathcal{X})} &\rightarrow 0 \quad \text{quand } k \rightarrow \infty, \\ \psi(\rho_{n_k}, x) - \psi(\hat{\rho}_x, x) &= \sum_j x^j \frac{\hat{\rho}_x^j - \rho_{n_k}^j}{s_j} - \frac{1}{2} (\|\rho_{n_k}\|_{\mathbf{u}}^2 - \|\hat{\rho}_x\|_{\mathbf{u}}^2). \end{aligned}$$

Mais

$$C = \int_0^T \Gamma(t, t) dt = \sum_j s_j > 0,$$

d'où $s_j < C$ pour tout j et par suite,

$$\sum_j \left(\frac{\rho^j}{s_j} \right)^2 = \|\rho\|_{\Gamma(\mathcal{X})}^2 > \frac{1}{C} \sum_j \frac{(\rho^j)^2}{s_j} = \frac{1}{C} \|\rho\|_{\mathbf{u}}^2 > \frac{1}{C^2} \sum_j (\rho^j)^2 = \frac{1}{C^2} \|\rho\|_{\mathcal{X}}^2$$

Nous avons donc

$$\lim_{k \rightarrow \infty} \psi(\rho_{n_k}, x) - \psi(\hat{\rho}_x, x) = 0$$

et conséquemment

$$\psi(\hat{\rho}_x, x) = M(x)$$

Si nous supposons que \mathcal{F} est un ensemble convexe, alors $\hat{\rho}_x$ vérifiant cette propriété est unique. En effet un calcul montre (voir [F₄]) que

$$\frac{1}{2} [\psi(\rho_1, x) + \psi(\rho_2, x)] = \psi\left(\frac{\rho_1 + \rho_2}{2}, x\right) - \frac{1}{8} \|\rho_1 - \rho_2\|_H^2,$$

et par suite

THEOREME 8. — *Si \mathcal{F} est un ensemble convexe et compact dans $\Gamma(\mathcal{X})$ et $f_\rho(x)$ est défini par (4.1) il existe une et une seule application $\hat{\rho}$ de \mathcal{X} dans \mathcal{F} telle que*

$$f_{\hat{\rho}_x}(x) \geq f_\rho(x) \quad \text{pour tout } \rho \in \mathcal{F}$$

Cette application nous définit un estimateur du maximum de vraisemblance

3. CAS STATIONNAIRE ET ESTIMATION D'UN DÉCALAGE DE LA MOYENNE. — Nous considérerons ici que la fonction aléatoire $U(t)$ est la partie relative à l'intervalle $[0, T]$ d'une fonction aléatoire stationnaire, de fonction de corrélation $r(h) = \Gamma(t, t + h)$.

\mathcal{F} sera un ensemble de fonctions dépendant du paramètre réel τ de la forme

$$\varphi_\tau(t) = \rho_0(t - \tau) \quad (\text{voir [H}_4]),$$

où ρ_0 est une fonction donnée et où τ est élément d'un intervalle également donné.

Alors le problème de l'estimation de la moyenne $E[X(t)]$ de $X(t)$ se réduit à l'estimation du nombre τ .

Plus précisément nous supposerons

1° que *a priori* τ peut prendre toutes les valeurs possibles de l'intervalle $[0, T - h]$

2° que ρ_0 est de la forme

$$\rho_0(t) = \int_0^h r(t - t') \lambda(t') dt' \quad \text{pour } -T + h \leq t \leq T,$$

où $\lambda(t)$ est une fonction de carré sommable sur $[0, h]$. On a pour $0 < \tau < T - h$,

$$\begin{aligned} \varphi_\tau(t) = \rho_0(t - \tau) &= \int_0^h r(t - \tau - t') \lambda(t') dt' \\ &= \int_{-\tau}^{\tau+h} r(t - \theta) \lambda(\theta - \tau) d\theta \quad \text{pour } 0 \leq t \leq T; \end{aligned}$$

donc, pour tout $0 < \tau < T - h$

$$\rho_\tau \in \Gamma(\mathcal{X})$$

et

$$\Gamma^{-1} \rho_\tau = \begin{cases} 0 & \text{pour } 0 \leq t \leq \tau, \\ \lambda(t - \tau) & \text{pour } \tau \leq t \leq \tau + h, \\ 0 & \text{pour } \tau + h < t \leq T. \end{cases}$$

Soit pour tout $0 \leq \tau \leq T - h$ et pour tout $x \in \mathcal{X}$,

$$f_\rho(x) = \exp \sum_j \frac{\rho_j^i}{s_j} \left(x_j - \frac{\rho_j^i}{2} \right) = \exp \left\{ \langle \Gamma^{-1} \rho_\tau, x \rangle - \frac{1}{2} \langle \Gamma^{-1} \rho_\tau, \rho_\tau \rangle \right\}.$$

Mais

$$\langle \Gamma^{-1} \rho_\tau, \rho_\tau \rangle = \int_\tau^{\tau+h} \rho_0(t - \tau) \lambda(t - \tau) dt = \int_0^h \rho_0(t) \lambda(t) dt;$$

donc, $\langle \Gamma^{-1} \rho_\tau, \rho_\tau \rangle$ est indépendant de τ .

Par suite, rendre maximal $f_\tau(x)$ (pour x fixé) revient à rendre maximal $\langle \Gamma^{-1} \rho_\tau, x \rangle$.

Or on a

$$\langle \Gamma^{-1} \rho_\tau, x \rangle = \int_\tau^{\tau+h} x(t) \lambda(t - \tau) dt$$

et si nous désignons $y(\alpha)$ la transformée de Fourier de $x(t)$ et par $\Lambda(\alpha)$ celle de $\lambda(t)$

$$\langle \Gamma^{-1} \rho_\tau, x \rangle = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\tau\alpha} \Lambda(\alpha) y(\alpha) d\alpha;$$

donc $\langle \Gamma^{-1} \rho_\tau, x \rangle$ est une fonction continue de τ sur l'intervalle $[0, T - h]$ et il existe $\hat{\tau}(x)$ tel que

$$\langle \Gamma^{-1} \rho_{\hat{\tau}}, x \rangle = \sup_{0 \leq \tau \leq T - h} \langle \Gamma^{-1} \rho_\tau, x \rangle.$$

THÉORÈME 9. — $\hat{\tau}$ est un estimateur du maximum de vraisemblance.

CHAPITRE III.

CERTAINS PROBLÈMES DE PRÉVISION.

Dans le présent chapitre, après quelques remarques générales sur la prévision à partir de l'observation d'une fonction aléatoire permanente sur un intervalle, nous étudierons deux problèmes de prévision particuliers.

1. Prévision à partir de l'observation sur un intervalle ⁽¹⁾ d'une fonction aléatoire permanente. — Soit (Ω, Σ, μ) un espace de probabilité. Et considérons une fonction aléatoire, c'est-à-dire une

(1) Pour le cas discret, voir [W₁].

famille $x_t(\omega)$ ($t \in \mathbb{R}$) d'applications mesurables de (Ω, Σ) dans (\mathbb{R}, ζ) [où ζ désigne le σ -corps des boréliens de la droite] et une variable aléatoire $y(\omega)$.

Examinons le problème suivant : à partir de l'observation de x_t pour $t \in T$, intervalle fini ou infini, estimer la valeur prise par y .

Si nous appliquons le critère des moindres carrés pour juger de la qualité de l'estimation \hat{y} de y , c'est-à-dire si nous voulons que \hat{y} soit telle que

$$E|\hat{y} - y|^2 = \int |\hat{y}(\omega) - y(\omega)|^2 d\mu(\omega)$$

soit le plus petit possible [pour que ceci ait un sens, il faut que les variables aléatoires introduites aient un second moment fini : nous supposons dans la suite que y a un second moment fini], on a le résultat suivant :

THÉOREME 1. — *Désignons par \mathcal{B}_t le σ -corps $x_t^{-1}(\zeta)$ et par $\bigcup_{t \in T} \mathcal{B}_t$ le σ -corps réunion de tous les \mathcal{B}_t pour tout $t \in T$.*

Alors $\min \int |y(\omega) - z(\omega)|^2 d\mu(\omega)$ pris pour tous les $z(\omega)$ mesurables par rapport à $\bigcup_{t \in T} \mathcal{B}_t$ est atteint pour

$$(1.1) \quad z_0(\omega) = E \left[y / \bigcup_{t \in T} \mathcal{B}_t \right].$$

En effet

$$\begin{aligned} & \int |y(\omega) - z(\omega)|^2 d\mu(\omega) \\ &= E |y(\omega) - z(\omega)|^2 = E |y(\omega) - z_0(\omega)|^2 \\ & \quad + E |z_0(\omega) - z(\omega)|^2 + 2 E [y(\omega) - z_0(\omega)] [z_0(\omega) - z(\omega)]. \end{aligned}$$

Mais $E[y(\omega) - z_0(\omega)] [z_0(\omega) - z(\omega)] = 0$, car si nous prenons d'abord l'espérance mathématique conditionnelle par rapport à $\bigcup_{t \in T} \mathcal{B}_t$,

on a

$$[z_0(\omega) - z(\omega)] E \left[y(\omega) - z_0(\omega) / \bigcup_{t \in T} \mathcal{B}_t \right] = 0,$$

d'où

$$E |y(\omega) - z(\omega)|^2 = E |y(\omega) - z_0(\omega)|^2 + E |z_0(\omega) - z(\omega)|^2$$

qui est minimal pour

$$z(\omega) = z_0(\omega) = E\left(y / \bigcup_{t \in T} \mathcal{B}_t\right).$$

THÉOREME 2. — *Il existe un ensemble dénombrable $(t_1, t_2, \dots, t_n, \dots)$ dense dans T tel que*

$$E\left(y / \bigcup_{t \in T} \mathcal{B}_t\right) = E\left(y / \bigcup_i \mathcal{B}_{t_i}\right) \quad \text{presque sûrement.}$$

En effet, on sait que si z est une fonction mesurable par rapport à $\bigcup_{t \in T} \mathcal{B}_t$, alors il existe un ensemble dénombrable $S \subset T$ (dépendant de z) tel que z est mesurable par rapport à $\bigcup_{t \in S} \mathcal{B}_t$ (voir [D₁] théorème 1.6, p. 604).

Donc, il existe un ensemble dénombrable $\{t_i\}$, $t_i \in T$, et une fonction $g(\omega)$ mesurable par rapport à $\bigcup_{i=1}^{\infty} \mathcal{B}_{t_i}$ telle que

$$E\left[y / \bigcup_{t \in T} \mathcal{B}_t\right] = g(\omega) \quad \text{presque partout.}$$

Prenons l'espérance mathématique conditionnelle par rapport à $\bigcup_i \mathcal{B}_{t_i}$, on a

$$g(\omega) = E\left[y / \bigcup_{i=1}^{\infty} \mathcal{B}_{t_i}\right] \quad \text{presque sûrement,}$$

d'où

$$E\left[y / \bigcup_{t \in T} \mathcal{B}_t\right] = E\left[y / \bigcup_i \mathcal{B}_{t_i}\right] \quad \text{presque sûrement.}$$

Ceci nous permet d'énoncer le théorème d'approximation suivant :

THÉOREME 3. — *Il existe un ensemble dénombrable t_i , ($t_i \in T$) et une suite de fonctions Φ_n telles que :*

1° Φ_n est mesurable de (R^n, ζ^n) dans (R, ζ) ;

2° $E\left[y / \bigcup_{t \in T} \mathcal{B}_t\right] = \lim_{n \rightarrow \infty} \Phi_n(x_{t_1}(\omega), \dots, x_{t_n}(\omega))$ presque sûrement.

En effet,

$$E\left(y/\bigcup_i \mathcal{B}_{t_i}\right) = \lim_{n \rightarrow \infty} E\left(y/\bigcup_{i=1}^n \mathcal{B}_{t_i}\right) \quad \text{presque sûrement.}$$

(voir [D₁] théorème 4.3, p. 331] et toute fonction mesurable par rapport à $\bigcup_{i=1}^n \mathcal{B}_{t_i}$ s'écrit $\Phi[x_{t_1}(\omega), \dots, x_{t_n}(\omega)]$, où Φ est une application mesurable de (R^n, ζ^n) dans (R, ζ) .

Les théorèmes précédents affirment l'existence d'une suite $\{t_i\}$ possédant les propriétés indiquées mais il serait utile d'avoir un moyen pour trouver une telle suite.

Dans le cas où x_t est un processus continu en probabilité, on peut prendre n'importe quel ensemble dénombrable dense dans T. Ceci résulte du théorème suivant (voir [A₁]).

THÉOREME 4. — Soit x_t un processus stochastique continu en probabilité et soit $\{t_i\}$ un ensemble dénombrable partout dense dans T.

Alors si $z(\omega)$ est une fonction mesurable par rapport à $\bigcup_{t \in T} \mathcal{B}_t$, il existe $z'(\omega)$ mesurable par rapport à $\bigcup_i \mathcal{B}_{t_i}$ telle que

$$z(\omega) = z'(\omega) \quad \text{presque sûrement.}$$

En effet, notons que $\bigcup_{t \in T} \mathcal{B}_t$ peut être construit de la façon suivante : on considère la classe \mathcal{D} des ensembles de la forme

$$B_{\tau_1} \cap B_{\tau_2} \cap \dots \cap B_{\tau_n} \quad \text{avec } B_{\tau_i} \in \mathcal{B}_{\tau_i} \quad (\tau_i \in T),$$

puis soit \mathcal{C} la classe des réunions finies d'éléments de \mathcal{D} (\mathcal{C} est un corps); alors $\bigcup_{t \in T} \mathcal{B}_t$ est le σ -corps engendré par \mathcal{C} . Soit $B_\tau \in \mathcal{B}_\tau$ ($\tau \in T$); on a $B_\tau = x_\tau^{-1}(E)$, $E \in \zeta$.

Mais il existe une suite $s_j \in \{t_i\}$ telle que

$$x_\tau(\omega) = \lim_{j \rightarrow \infty} x_{s_j}(\omega) \quad \text{presque sûrement.}$$

Alors si l'on pose

$$\begin{aligned} x'_\tau(\omega) &= \lim x_{s_j}(\omega) && \text{lorsque cette limite existe,} \\ &= 0 && \text{dans le cas contraire,} \end{aligned}$$

$x'_\tau(\omega)$ est mesurable $\bigcup_i \mathcal{B}_{t_i}$ et $x'_\tau(\omega) = x_\tau(\omega)$ presque sûrement.

Soit alors

$$B'_\tau = x'^{-1}_\tau(E).$$

On a

$$\mu(B_{\tau\Delta} B'_\tau) = 0$$

(où Δ désigne la différence symétrique). Pour

$$D = B_{\tau_1} \cap \dots \cap B_{\tau_n},$$

posons

$$D' = B'_{\tau_1} \cap \dots \cap B'_{\tau_n}.$$

On a

$$\mu(D_\Delta D') = 0, \quad D' \in \bigcup_i \mathcal{B}_{t_i}.$$

Pour

$$\Lambda = D_1 \cup \dots \cup D_m \in \mathcal{C}, \quad \text{soit } \Lambda' = D'_1 \cup \dots \cup D'_m,$$

on a

$$\mu(\Lambda_\Delta \Lambda') = 0 \quad \text{et} \quad \Lambda' \in \bigcup_i \mathcal{B}_{t_i}.$$

Soit alors K la classe des $\Lambda \in \bigcup_{t \in T} \mathcal{B}_t$ tels que pour tout $\Lambda \in K$, il

$$\Lambda' \in \bigcup_i \mathcal{B}_{t_i} \quad \text{avec} \quad \mu(\Lambda_\Delta \Lambda') = 0,$$

existe c'est une classe monotone et elle contient le corps \mathcal{C} , donc elle est égale à $\bigcup_{t \in T} \mathcal{B}_t$.

Par suite ([D₁] théorème 2.3, p. 605) on a le théorème énoncé.

Après ces quelques considérations générales, nous allons étudier deux problèmes particuliers de prévision.

2. Un problème de détermination d'un filtrage optimal. — Le problème suivant de filtrage en présence d'un bruit laplacien, pour certains types de messages a été étudié par G. Kallianpur (voir [K₁]).

Nous introduirons ici une démonstration différente, utilisant les densités de probabilités entre processus laplaciens, considérées dans les chapitres précédents. Ceci nous permettra en même temps d'élargir les conditions imposées par Kallianpur.

(Ω, \mathcal{O}, p) désignera l'espace des épreuves.

Considérons une fonction aléatoire laplacienne, réelle, mesurable, de moyenne nulle et de covariance $\Gamma(t, \tau)$ continue et connue; soit $U_t(\omega)$.

Soit d'autre part,

$$(2.1) \quad M_t(\omega) = \sum_{k=1}^n M_k(\omega) g_k(t),$$

où les M_k sont des variables aléatoires, réelles, indépendantes de U_t et de fonction de répartition conjointe connue : $F(m_1, m_2, \dots, m_n)$; et où les fonctions $g_k(t)$ sont des fonctions certaines, réelles, connues sous lesquelles nous ne ferons pour l'instant que l'hypothèse qu'elles sont continues.

Considérons alors la question suivante : ayant observé

$$X_t(\omega) = M_t(\omega) + U_t(\omega) \quad \text{pour } 0 \leq t \leq T,$$

que nous apporte cette observation sur la connaissance des M_k ?

Nous obtiendrons la loi de probabilité conjointe des M_k conditionnellement lorsque $X_t(\omega)$, $0 \leq t \leq T$ est fixé. Il est alors aisé de déduire la meilleure (au sens des moindres carrés) prévision de $M_0(\omega)$ à partir de l'observation de $X_t(\omega)$ sur l'intervalle $[0, T]$ (ce qui est le sujet de $[K_1]$).

Comme nous l'avons vu précédemment, $U_t(\omega)$ peut être considéré comme un élément aléatoire $U(\omega)$ à valeurs dans \mathcal{X} , espace de Hilbert des fonctions réelles, définies et de carré sommable sur $[0, T]$; si l'on ne considère que la loi temporelle, la mesure correspondante μ_0 est définie sur \mathcal{A} , plus petit σ -corps de sous-ensemble de \mathcal{X} par rapport auquel les applications h_t qui à tout $x \in \mathcal{X}$ associe sa valeur au point $t \in [0, T]$ sont mesurables; $\mathcal{A} = \bigcup_{t \in [0, T]} h_t^{-1}(\zeta)$, où ζ est le σ -corps des boréliens de la droite.

De même, d'après les hypothèses faites, $X_t(\omega)$ peut être considéré comme un élément aléatoire $X(\omega)$ à valeurs dans \mathcal{X} ; soit ν la mesure correspondante définie sur \mathcal{A} .

Désignons par \mathfrak{T} , l'application de Ω dans \mathfrak{X} qui à ω fait correspondre $X(\omega)$.

\mathfrak{T} est une application mesurable de (Ω, \mathcal{D}) dans $(\mathfrak{X}, \mathcal{A})$ et plus précisément on a, comme on le vérifie aisément,

$$(2.2) \quad \bigcup_{t \in [0, T]} X_t^{-1}(\zeta) = \mathfrak{T}^{-1}(\mathcal{A})$$

(\bigcup désignant réunion de σ -corps).

Remarquons aussi que

$$\nu(A) = p(\mathfrak{T}^{-1}(A)) \quad \text{pour } A \in \mathcal{A};$$

ce que nous noterons, comme il est coutume (voir [H₁]) par $\nu = p\mathfrak{T}^{-1}$.

Soit M l'application qui à ω associe $M_1(\omega), \dots, M_n(\omega)$.

Considérons alors la probabilité conditionnelle lorsque M est fixé; soit $p_m(\cdot)$.

Soit $\mu_m = p_m \mathfrak{T}^{-1}$; μ_m est la mesure de probabilité correspondante à l'élément aléatoire $\sum_{k=1}^n m_k g_k + U(\omega)$.

En reprenant les notations du paragraphe 1, du chapitre I, nous savons (théorème 2, chapitre I) que les mesures μ_m constituent un ensemble de mesures équivalentes si $g_k \in H$, pour tout $k = 1, \dots, n$, ce que nous supposerons dans la suite.

Désignons par $f(x, m)$ la densité de probabilité de μ_m par rapport à μ_0 . Si p_m^* désigne la restriction de p_m à $\mathfrak{T}^{-1}(\mathcal{A})$, alors p_m^* est absolument continue par rapport à p_0^* et

$$(2.3) \quad \frac{d p_m^*}{d p_0^*}(\omega) = f(\mathfrak{T}(\omega), m).$$

De l'expression de $f(x, m)$ on déduit aisément que $(\omega, m) \rightarrow f(\mathfrak{T}(\omega), m)$ est une application mesurable de $(\Omega \times \mathbb{R}^n, \mathfrak{T}^{-1}(\mathcal{A}) \times \zeta^n)$ dans (\mathbb{R}, ζ) .

On a pour $B \in \mathfrak{T}^{-1}(\mathcal{A})$,

$$p(B) = \int_{\Omega} p_m(B) p(d\omega) = \int_{\mathbb{R}^n} p_m(B) F(dm) = \int_{\mathbb{R}^n} F(dm) \int_B f(\mathfrak{T}(\omega), m) p_0(d\omega).$$

En appliquant le théorème de Fubini,

$$(2.4) \quad = \int_B p_0(d\omega) \int_{\mathbb{R}^n} f(\mathfrak{T}(\omega), m) F(dm).$$

Soit $E_n \in \mathcal{Z}^n$; la probabilité conditionnelle pour que $M \in E_n$ lorsque X_t , $0 \leq t \leq T$, est fixé, est par définition

$$\Pr\left(M \in E_n / \bigcup_{t \in [0, T]} X_t^{-1}(\zeta)\right) = \Pr[M \in E_n / \mathfrak{T}^{-1}(\mathcal{A})];$$

c'est une fonction mesurable en ω par rapport à $\mathfrak{T}^{-1}(\mathcal{A})$ qui doit vérifier

$$p(M^{-1}(E_n) \cap B) = \int_B \Pr[M \in E_n / \mathfrak{T}^{-1}(\mathcal{A})] p(d\omega) \quad \text{pour tout } B \in \mathfrak{T}^{-1}(\mathcal{A}).$$

Mais

$$\begin{aligned} p(M^{-1}(E_n) \cap B) &= \int_{M^{-1}(E_n)} p_m(B) p(d\omega) = \int_{E_n} p_m(B) F(dm) \\ &= \int_{E_n} F(dm) \int_B f(\mathfrak{T}(\omega), m) p_0(d\omega) \\ &= \int_B p_0(d\omega) \int_{E_n} f(\mathfrak{T}(\omega), m) F(dm) \\ &= \int_B \frac{\int_{E_n} f(\mathfrak{T}(\omega), m) F(dm)}{\int_{\mathbb{R}^n} f(\mathfrak{T}(\omega), m) F(dm)} p_0(d\omega) \int_{\mathbb{R}^n} f(\mathfrak{T}(\omega), m) F(dm), \end{aligned}$$

d'où d'après (2.4),

$$(2.5) \quad = \int_B \frac{\int_{E_n} f(\mathfrak{T}(\omega), m) F(dm)}{\int_{\mathbb{R}^n} f(\mathfrak{T}(\omega), m) F(dm)} p(d\omega)$$

La fonction $\frac{\int_{E_n} f(\mathfrak{T}(\omega), m) F(dm)}{\int_{\mathbb{R}^n} f(\mathfrak{T}(\omega), m) F(dm)}$ est une fonction mesurable en ω

par rapport à $\mathfrak{T}^{-1}(\mathcal{A})$; et par suite

THÉOREME 5. — *La probabilité conditionnelle du vecteur aléatoire M , étant donné X_t , $0 \leq t \leq T$, s'écrit*

$$(2.6) \quad \Pr(M \in E_n / X_t, 0 \leq t \leq T) = \frac{\int_{E_n} f(\mathfrak{T}(\omega), m) F(dm)}{\int_{\mathbb{R}^n} f(\mathfrak{T}(\omega), m) F(dm)}.$$

Supposons qu'à partir de l'observation de $X_t(\omega)$ sur l'intervalle $0 \leq t \leq T$, nous désirions construire la « meilleure » prévision de $M_0(\omega)$, meilleur étant entendu au sens des moindres carrés.

Nous savons (théorème 1) que parmi toutes les fonctions mesurables par rapport à $\bigcup_{0 \leq t \leq T} X_t^{-1}(\zeta)$, celle qui minimise $\int |Z(\omega) - M_0(\omega)|^2 \rho(d\omega)$ est la moyenne conditionnelle de $M_0(\omega)$ lorsque $X_t(\omega)$ pour tout $0 \leq t \leq T$, est fixé :

$$(2.7) \quad E(M_0(\omega)/X_t, 0 \leq t \leq T) \\ = \frac{1}{\int_{\mathbb{R}^n} f(\mathfrak{Z}(\omega), m) F(dm)} \sum_{k=1}^n g_k(\theta) \int_{\mathbb{R}^n} m_k f(\mathfrak{Z}(\omega), m) F(dm).$$

Remarque. — Le problème de l'expression de $f(x, m)$ ainsi que celui de la vérification des conditions imposées sur les fonctions $g_k(t)$, (c'est-à-dire appartenance à H), se posent pour une utilisation pratique des formules :

1° Dans le cas où les g_k sont des éléments de l'espace $\Gamma(\mathfrak{X})$ [alors nécessairement ils sont dans H, puisque $\Gamma(\mathfrak{X}) \subset H$]; comme nous l'avons vu au chapitre II, $f(x, m)$ a alors une expression simple.

C'est dans ce cas que Kallianpur a obtenu la formule (2.7) en supposant de plus que F admettait une densité.

2° Mais comme nous l'avons vu également au chapitre II, pour des g_k appartenant à H mais non nécessairement à $\Gamma(\mathfrak{X})$, des cas simples et importants où $f(x, m)$ a une expression simple en fonction des données, nous sont fournis par le processus de Wiener-Lévy, et par celui de Ornstein-Uhlenbeck ou plus généralement par les processus dits auto-régressifs.

Remarquons également que l'étude précédente nous fournit un exemple d'une meilleure prévision (au sens des moindres carrés) qui n'est pas nécessairement linéaire.

3. Prévision linéaire dans le cas de fonctions aléatoires dont les moyennes sont des combinaisons linéaires inconnues de fonctions certaines connues. — Dans ce paragraphe, nous ne considérerons cette fois que les prévisions linéaires. Soient

$$(3.1) \quad \begin{aligned} X(t) &= m(t) + X_0(t) \\ Y(t) &= m(t) + Y_0(t) \end{aligned} \quad (t \in \mathbb{R}),$$

où $X_0(t)$ et $Y_0(t)$ sont des fonctions aléatoires du second ordre, dont

les moyennes sont identiquement nulles et dont les covariances ainsi que la fonction $E[Y_0(t) \overline{X_0(\tau)}]$ sont connues ; et où

$$(3.2) \quad m(t) = \sum_{k=1}^n a_k g_k(t),$$

avec $g_k(t)$ fonctions certaines, connues et a_k coefficients certains mais inconnus.

Supposons qu'à partir de l'observation de $X(t)$ sur un intervalle fini que nous pouvons toujours poser égal à $[-T, 0]$, on désire prévoir la valeur prise par $Y(s)$ à l'instant $s > 0$.

Par exemple, $m(t)$ pourrait représenter une coordonnée d'un mobile et $X(t)$ la mesure de cette coordonnée par un système qui reçoit de la part du mobile des données en présence de bruit ($X_0(t)$). Si l'on désirait prévoir la position du mobile à une date future, alors nous serions dans le cas du problème posé avec $Y_0(t) \equiv 0$.

Nous nous limiterons aux prévisions linéaires et nous utiliserons pour juger la qualité d'une prévision, le critère des moindres carrés. Plus précisément nous nous proposons de chercher la prévision linéaire non biaisée de $Y(s)$, la meilleure au sens des moindres carrés.

Soit $\mathcal{K}_X(-T, 0)$ le sous-espace de l'espace de Hilbert \mathcal{K} des variables aléatoires ayant un second moment fini (\mathcal{K} est muni du produit scalaire habituel $\langle \xi, \eta \rangle = E\xi\bar{\eta}$) engendré par $X(t)$ pour $-T \leq t \leq 0$.

Une prévision linéaire non biaisée de $Y(s)$ sera un élément $W \in \mathcal{K}_X(-T, 0)$ tel que $E[W] = m(s)$ quels que soient a_1, \dots, a_n .

Nous cherchons parmi de tels W à rendre minimal $E|W - Y(s)|^2$. Supposons qu'il existe un tel W qui réalise le minimum : soit W_0 . Alors si l'on cherche une meilleure estimation non biaisée $\in \mathcal{K}_X(-T, 0)$ d'une quantité de la forme

$$Y(s) + \sum_{j=1}^m b_j X(-t_j), \quad \text{avec } 0 < t_j < T,$$

il est clair que $W_0 + \sum_{j=1}^m b_j X(-t_j)$ en est une.

Réciproquement, si W_1 est une meilleure estimation non biaisée $\in \mathcal{K}_X(-T, 0)$ de $Y(s) + \sum_{j=1}^m b_j X(-t_j)$, avec $0 < t_j < T$, alors

$W_1 - \sum_{j=1}^m b_j X(-t_j)$ est une meilleure estimation non biaisée $\in \mathcal{K}_X(-T, 0)$ de $Y(s)$.

Ceci étant remarqué, on peut se ramener à l'estimation d'une variable aléatoire ayant une moyenne nulle.

Si nous supposons que les fonctions $g_k(t)$ sont linéairement indépendantes sur $(-T, 0)$, alors il existe n points $0 < t_1 < t_2 < \dots < t_n < T$ tels que le déterminant $[g_k(-t_j)]_{k,j}$ soit $\neq 0$. Considérons la combinaison linéaire

$$(3.3) \quad m(s) + P_1 m(-t_1) + \dots + P_n m(-t_n) \equiv 0$$

quels que soient $a_1, a_2, \dots, a_n, P_1, P_2, \dots, P_n$ sont donnés par les équations

$$(3.4) \quad g_k(s) + \sum_{j=1}^n g_k(-t_j) P_j = 0 \quad (k = 1, \dots, n).$$

Si nous posons

$$(3.5) \quad Z = Y(s) + P_1 X(-t_1) + \dots + P_n X(-t_n)$$

Alors

$$E[Z] = 0 \quad \text{quels que soient } a_1, \dots, a_n.$$

Nous sommes ramenés à trouver parmi les fonctionnelles linéaires sur les $X(t)$ ($-T \leq t \leq 0$), annulant les fonctions de la forme

$$m(t) = \sum_{j=1}^n a_j g_j(t) \text{ une } \tilde{Z} \text{ pour laquelle } E|Z - \tilde{Z}|^2 \text{ soit minimal.}$$

Les éléments $W \in \mathcal{K}_X(-T, 0)$ tels que $E[W] = 0$ quels que soient a_1, \dots, a_n sont de la forme

$$\sum_{i=1}^m \alpha_i X(\tau_i),$$

avec

$$-T \leq \tau_i \leq 0 \quad \text{et} \quad \sum_{i=1}^m \alpha_i g_k(\tau_i) = 0 \quad \text{pour } k = 1, \dots, n.$$

ou limite en moyenne quadratique quand $j \rightarrow \infty$ de

$$\sum_{i=1}^{m_j} \alpha'_i X(\tau'_i),$$

avec

$$-T \leq \tau'_k \leq 0 \quad \text{et} \quad \sum_{i=1}^{m_k} \alpha'_i g_k(\tau'_i) \rightarrow 0 \quad \text{pour} \quad k = 1, \dots, n.$$

Soit E l'ensemble de ces éléments.

On constate aisément que dans \mathcal{K} , toute limite d'éléments de E appartient à E et que, quels que soient U et V appartenant à E et $0 < \mu < 1$, $\mu U + (1 - \mu) V$ est encore élément de E.

Autrement dit E est un ensemble fermé et convexe de \mathcal{K} . Par suite, il existe un et un seul élément $\tilde{Z} \in E$ tel que

$$(3.6) \quad E | Z - \tilde{Z} |^2 = \min_{U \in E} E | Z - U |^2.$$

D'où :

THÉOREME 6. — *Il existe une meilleure prévision linéaire non biaisée de Y(s) à partir de X(t) ($-T \leq t \leq 0$) et elle est unique.*

En effet, c'est

$$(3.7) \quad \tilde{Y}(s) = \tilde{Z} - \sum_{i=1}^n P_i X(-t_i)$$

THÉOREME 7. — *Il existe des coefficients c_1, \dots, c_n tels que \tilde{Y} vérifie l'équation*

$$(3.8) \quad E[Y(s) - \tilde{Y}(s)] \overline{X(t)} = \sum_{k=1}^n c_k g_k(t) \quad \text{pour tout} \quad -T \leq t \leq 0.$$

Et cette équation jointe à la condition $\tilde{Y}(s)$ est une prévision linéaire non biaisée de Y(s) à partir de X(t) ($-T \leq t \leq 0$) détermine complètement $\tilde{Y}(s)$.

Soient $t_1, \dots, t_n \in [-T, 0]$ tels que le déterminant $[g_k(t_j)]_{k,j}$ soit différent de zéro. Définissons les coefficients c_1, \dots, c_n par le système d'équations linéaires

$$E[Z - \tilde{Z}] \overline{X(t)} = \sum_{k=1}^n c_k g_k(t) \quad (t = t_1, \dots, t_n).$$

Soit $t \in [-T, 0]$ et différent de t_1, \dots, t_n .

Désignons par $L(t_1, \dots, t_n, t)$ la combinaison linéaire

$$\sum_{j=1}^n b_j X(t_j) + X(t)$$

telle que

$$\sum_{j=1}^n b_j g_k(t_j) + g_k(t) = 0 \quad (k=1, \dots, n).$$

Alors ε étant un nombre complexe quelconque, on a

$$\tilde{Z}_1 = \tilde{Z} + \varepsilon L(t_1, \dots, t_n, t) \in E.$$

Calculons

$$\begin{aligned} E|Z - \tilde{Z}_1|^2 &= E|Z - \tilde{Z}|^2 + |\varepsilon|^2 E|L(t_1, \dots, t_n, t)|^2 \\ &\quad - 2\Re\{\varepsilon E[Z - \tilde{Z}] \overline{L(t_1, \dots, t_n, t)}\}. \end{aligned}$$

Or d'après la propriété de minimum de $E|Z - \tilde{Z}|^2$, on doit avoir

$$E|Z - \tilde{Z}|^2 \leq E|Z - \tilde{Z}_1|^2 \quad \text{quel que soit } \varepsilon.$$

Ce qui implique que

$$E[Z - \tilde{Z}] \overline{L(t_1, \dots, t_n, t)} = 0,$$

c'est-à-dire

$$\begin{aligned} \sum_{j=1}^n b_j \sum_{k=1}^n c_k g_k(t_j) + E[Z - \tilde{Z}] \overline{X(t)} &= 0, \\ E[Z - \tilde{Z}] \overline{X(t)} &= - \sum_{k=1}^n c_k \sum_{j=1}^n b_j g_k(t_j) = \sum_{k=1}^n c_k g_k(t). \end{aligned}$$

L'équation

$$E[Z - \tilde{Z}] \overline{X(t)} = \sum_{k=1}^n c_k g_k(t) \quad \text{pour } -T \leq t \leq 0$$

jointe à la condition $\tilde{Z} \in E$ détermine complètement \tilde{Z} .

Soit en effet $\tilde{Z}_1 \in E$ tel que

$$E[Z - \tilde{Z}_1] \overline{X(t)} = \sum_{k=1}^n c'_k g_k(t) \quad (-T \leq t \leq 0).$$

$$\tilde{Z} = \lim_{j \rightarrow \infty} (\text{en m.q.}) \sum_{i=1}^{m_j} \alpha_i^j X(\tau_i^j),$$

$$\begin{aligned} E[Z - \tilde{Z}_1] \overline{\tilde{Z}} &= \lim_{j \rightarrow \infty} E[Z - \tilde{Z}_1] \sum_{i=1}^{m_j} \alpha_i^j \overline{X(\tau_i^j)} \\ &= \lim_{j \rightarrow \infty} \sum_{i=1}^{m_j} \alpha_i^j \sum_{k=1}^n c_k g_k(\tau_i^j) \\ &= \lim_{j \rightarrow \infty} \sum_{k=1}^n c_k \sum_{i=1}^{m_j} \alpha_i^j g_k(\tau_i^j) = 0. \end{aligned}$$

De même,

$$E[Z - \tilde{Z}_1] \overline{\tilde{Z}_1} = 0,$$

d'où

$$E[Z - \tilde{Z}_1] [\overline{\tilde{Z} - \tilde{Z}_1}] = 0,$$

et aussi

$$E[Z - \tilde{Z}] [\overline{\tilde{Z} - \tilde{Z}_1}] = 0;$$

par suite,

$$E|Z - Z_1|^2 = 0.$$

Remarque. — Si $g_k(t) \equiv 0$ pour tout k , l'équation (3,8) devient évidemment l'équation de la projection de $Y(s)$ sur $\mathcal{K}_X(-T, 0)$.

Nous allons voir maintenant comment les méthodes introduites par Yaglom (*voir* [Y_1]) permettent de résoudre l'équation (3.8) dans le cas où :

1° Les fonctions aléatoires $X_0(t)$ et $Y_0(t)$ sont stationnaires, stationnairement corrélées entre elles et ont des densités spectrales de corrélation et d'intercorrélation rationnelles.

2° Les $g_k(t)$ sont des fonctions continues sur $[-T, 0]$ et $\frac{1}{2\pi} \int_{-T}^0 e^{i\lambda t} m(t) dt$ peut se mettre sous la forme

$$M^{(1)}(\lambda) + e^{-i\lambda T} M^{(2)}(\lambda),$$

où $M^{(1)}(\lambda)$ et $M^{(2)}(\lambda)$ sont des fonctions méromorphes ayant p pôles et telles que

$$\begin{aligned} M^{(1)}(\lambda) &\rightarrow 0 \quad \text{uniformément quand } |\lambda| \rightarrow \infty \\ &\quad \text{dans le demi-plan supérieur;} \\ M^{(2)}(\lambda) &\rightarrow 0 \quad \text{uniformément quand } |\lambda| \rightarrow \infty \\ &\quad \text{dans le demi-plan inférieur.} \end{aligned}$$

Remarquons que le (2°) est vérifié en particulier dans les deux cas suivants :

$$\begin{aligned} (3.9) \quad A & \quad g_k(t) = t^{k-1}, \\ (3.10) \quad B & \quad g_k(t) = e^{i\lambda_k t}; \end{aligned}$$

c'est-à-dire dans le cas où $m(t)$ est un polynome de degré $(n-1)$ en t et celui où $m(t)$ est un polynome trigonométrique.

Remarquons également que dans le cas où $g_k(t) = t^{k-1}$, $X(t)$ est une fonction aléatoire à $n^{\text{ième}}$ accroissements stationnaires (*voir* [Y_2]).

Dans cet article, Yaglom a considéré aussi l'estimation de quantités telles que Z à partir de $X(t)$, $-T \leq t \leq 0$.

A cette fin, il projette Z sur $\mathcal{K}_{\Delta_\tau^n X}(-T, 0)$, sous-espace de \mathcal{K} engendré par $\Delta_\tau^n X(t)$ pour $-T < t \leq 0$, $0 < \tau < \frac{T-t}{n}$ [Δ_τ^n désignant la $n^{\text{ième}}$ itérée de Δ_τ définie par $\Delta_\tau X(t) = X(t) - X(t - \tau)$]. Il ne semble pas qu'en général $\mathcal{K}_{\Delta_\tau^n X}(-T, 0) = E$; toutefois dans le cas présent les résultats montreront que les projections de Z sur E et sur $\mathcal{K}_{\Delta_\tau^n X}(-T, 0)$ sont confondues.

Soient

$$(3.11) \quad E[X_0(t+h)\bar{X}_0(t)] = \int_{-\infty}^{+\infty} e^{it\lambda} f(\lambda) d\lambda,$$

$$(3.12) \quad E[Y_0(t+h)\bar{X}_0(t)] = \int_{-\infty}^{+\infty} e^{it\lambda} f_1(\lambda) d\lambda.$$

On sait que \mathcal{K}_{X_0} sous-espace de \mathcal{K} , engendré par $X_0(t)$ ($-\infty < t < +\infty$) est en correspondance isométrique avec \mathcal{L}_F^2 espace des fonctions du carré de module sommable par rapport à $f(\lambda) d\lambda$.

Dans cette isométrie $e^{i\lambda t}$ correspond à $X_0(t)$ et si la décomposition spectrale du processus $X_0(t)$ lui-même s'écrit

$$X_0(t) = \int_{-\infty}^{+\infty} e^{i\lambda t} d\xi(\lambda),$$

où $\xi(\lambda)$ est une fonction aléatoire à accroissements non-corrélés telle que

$$E|\xi(\lambda + \Delta\lambda) - \xi(\lambda)|^2 = \int_{\lambda}^{\lambda + \Delta\lambda} f(\lambda) d\lambda,$$

alors l'isométrie entre \mathcal{K}_{X_0} et \mathcal{L}_F^2 peut être décrite par

$$(3.13) \quad \varphi(\lambda) \leftrightarrow \int_{-\infty}^{+\infty} \varphi(\lambda) d\xi(\lambda),$$

$\tilde{Z} \in \mathcal{K}_{X_0}(-T, 0)$ et s'écrit donc

$$(3.14) \quad \tilde{Z} = \int_{-\infty}^{+\infty} \varphi_s(\lambda) d\xi(\lambda),$$

où $\varphi_s(\lambda)$ appartient au sous-espace de \mathcal{L}_F^2 engendré par $e^{i\lambda t}$ pour $-T \leq t \leq 0$.

Remarquons que la relation (3.8), par suite de la continuité en moyenne quadratique de $X_0(t)$ et de la continuité de $g_k(t)$, est équivalente à la relation (3.8').

$$(3.8') \quad E[Z - \tilde{Z}] \bar{X}(t) = \sum_{k=1}^n c_k g_k(t) \quad (-T < t < 0),$$

d'où $\varphi_s(\lambda)$ doit vérifier

$$\begin{aligned} \int_{-\infty}^{+\infty} e^{i\lambda(s-t)} f_1(\lambda) d\lambda + \int_{-\infty}^{+\infty} \left[\sum_{j=1}^n P_j e^{-i\lambda t_j} - \varphi_s(\lambda) \right] e^{-i\lambda t} f(\lambda) d\lambda \\ = \sum_{k=1}^n c_k g_k(t) \quad (-T < t < 0), \end{aligned}$$

c'est-à-dire

$$(3.15) \quad \int_{-\infty}^{+\infty} \left\{ e^{i\lambda s} f_1(\lambda) + \left[\sum_{j=1}^n P_j e^{-i\lambda t_j} - \varphi_s(\lambda) \right] f(\lambda) - M^{(1)}(\lambda) - e^{-i\lambda T} M^{(2)}(\lambda) \right\} e^{-i\lambda t} d\lambda = 0 \quad \text{pour } -T < t < 0.$$

Donc nous sommes ramenés à trouver une fonction $\varphi_s(\lambda)$ qui vérifie les conditions suivantes :

- a. $\varphi_s(\lambda) \in \mathcal{L}_F^2(-T, 0)$;
- b. dans l'isométrie entre L_F^2 et \mathcal{H}_{X_0} , $\varphi(\lambda) \leftrightarrow \int_{-\infty}^{+\infty} \varphi(\lambda) d\zeta(\lambda)$, il correspond à $\varphi_s(\lambda)$ une fonctionnelle linéaire sur $X_0(t)$ qui s'annule si l'on remplace $X_0(t)$ par $g_k(t)$ ($k = 1, \dots, n$);

c. on a

$$(3.16) \quad \int_{-\infty}^{+\infty} \left\{ e^{i\lambda s} f_1(\lambda) + \left[\sum_{j=1}^n P_j e^{-i\lambda t_j} - \varphi_s(\lambda) \right] f(\lambda) - M^{(1)}(\lambda) - e^{-i\lambda T} M^{(2)}(\lambda) \right\} e^{i\lambda t} d\lambda = 0 \quad \text{pour } 0 < t < T.$$

On peut remarquer que

$$\int_{-\infty}^{+\infty} \psi_1(\lambda) e^{i\lambda t} d\lambda = 0 \quad \text{pour } t > 0,$$

si $\psi_1(\lambda)$ est une fonction prolongeable en une fonction holomorphe dans le demi-plan supérieur et tendant uniformément vers zéro quand $|\lambda| \rightarrow \infty$ dans ce demi-plan.

Et que

$$\int_{-\infty}^{+\infty} \psi_2(\lambda) e^{i\lambda t} d\lambda = 0 \quad \text{pour } t < 0,$$

si $\psi_2(\lambda)$ est une fonction prolongeable en une fonction holomorphe dans le demi-plan inférieur et tendant uniformément vers zéro quand $|\lambda| \rightarrow \infty$ dans ce demi-plan.

On constate alors que

$$\int_{-\infty}^{+\infty} \psi(\lambda) e^{i\lambda t} d\lambda = 0 \quad \text{pour } 0 < t < T,$$

si $\psi(\lambda)$ peut se mettre sous la forme $\psi_1(\lambda) + e^{-i\lambda T} \psi_2(\lambda)$, où $\psi_1(\lambda)$ et $\psi_2(\lambda)$ sont prolongeables en des fonctions holomorphes respectivement dans le demi-plan supérieur et dans le demi-plan inférieur et tendant uniformément vers 0 quand $|\lambda| \rightarrow \infty$ dans leur demi-plan respectif.

De la même façon, on voit aisément que si l'on prend pour $\varphi_s(\lambda)$ une fonction prolongeable en une fonction entière de la forme

$$(3.17) \quad \varphi_s(\lambda) = \Phi^{(1)}(\lambda) + e^{-i\lambda T} \Phi^{(2)}(\lambda) + \sum_{j=1}^n P_j e^{-i\lambda t_j},$$

avec $\Phi^{(1)}(\lambda)$ et $\Phi^{(2)}(\lambda)$ fractions rationnelles appartenant à \mathcal{L}_F^2 alors on peut affirmer que $\varphi_s(\lambda) \in \mathcal{L}_F^2(-T, 0)$.

Considérons une fonction $\varphi_s(\lambda)$ de cette forme.

En faisant apparaître les polynômes dans la décomposition de $\Phi^{(1)}(\lambda)$ et $\Phi^{(2)}(\lambda)$, on peut écrire

$$(3.18) \quad \Phi^{(1)}(\lambda) + e^{-i\lambda T} \Phi^{(2)}(\lambda) = \sum_k (\omega_k + \omega'_k e^{-i\lambda T}) (i\lambda)^k + \varphi^*(\lambda),$$

où $\varphi^*(\lambda)$ est de carré sommable par rapport à la mesure de Lebesgue sur $(-\infty, +\infty)$.

Si $\nu(t)$ est telle que

$$(3.19) \quad \varphi^*(\lambda) = \int_{-T}^0 e^{i\lambda t} \nu(t) dt,$$

On a

$$(3.20) \quad \int_{-\infty}^{+\infty} \varphi_s(\lambda) d\zeta(\lambda) = \sum_{j=1}^n P_j X_0(-t_j) + \sum_k [\omega_k X_0^{(k)}(0) + \omega'_k X_0^{(k)}(-T)] + \int_{-T}^0 X_0(t) \nu(t) dt,$$

où $X_0^{(k)}(t)$ désigne la dérivée d'ordre k en moyenne quadratique.

La condition (b) sera vérifiée si

$$(3.21) \quad \sum_{j=1}^n P_j g_l(-t_j) + \sum_k [\omega_k g_l^{(k)}(0) + \omega'_k g_l^{(k)}(-T)] + \int_{-T}^0 g_l(t) v(t) dt = 0$$

pour $l = 1, \dots, n$.

Quand à la relation (3.16) elle sera certainement vérifiée si

$$(3.22) \quad \begin{aligned} \psi_1(\lambda) &= e^{i\lambda s} f_1(\lambda) - \Phi^{(1)}(\lambda) f(\lambda) - M^{(1)}(\lambda), \\ \psi_2(\lambda) &= -\Phi^{(2)}(\lambda) f(\lambda) - M^{(2)}(\lambda) \end{aligned}$$

sont prolongeables en des fonctions holomorphes respectivement dans le demi-plan supérieur et dans le demi-plan inférieur et tendant uniformément vers zéro quand $|\lambda| \rightarrow \infty$ dans leur demi-plan respectif.

Soient

$$(3.23) \quad f(\lambda) = \frac{(\lambda - \beta_1) \dots (\lambda - \beta_M) (\lambda - \bar{\beta}_1) \dots (\lambda - \bar{\beta}_M)}{(\lambda - \alpha_1) \dots (\lambda - \alpha_N) (\lambda - \bar{\alpha}_1) \dots (\lambda - \bar{\alpha}_N)},$$

les parties imaginaires des α_i et des β_j étant positives et

$$(3.24) \quad f_1(\lambda) = \frac{D(\lambda)}{(\lambda - \gamma_1) \dots (\lambda - \gamma_{L_1}) (\lambda - \gamma'_1) \dots (\lambda - \gamma'_{L_2})},$$

où les parties imaginaires des γ_i sont positives et celles des γ'_j négatives.

Soient également $\delta_1, \dots, \delta_p$ les pôles nécessairement communs de $M^{(1)}(\lambda)$ et $M^{(2)}(\lambda)$ [ces deux fonctions ont les mêmes pôles puisque $M^{(1)}(\lambda) + e^{-i\lambda T} M^{(2)}(\lambda)$ est une fonction entière].

Enfin nous poserons pour alléger l'écriture,

$$\begin{aligned} (\lambda - \alpha_1) \dots (\lambda - \alpha_N) &= A(\lambda), & (\lambda - \bar{\alpha}_1) \dots (\lambda - \bar{\alpha}_N) &= \bar{A}(\lambda), \\ (\lambda - \beta_1) \dots (\lambda - \beta_M) &= B(\lambda), & (\lambda - \bar{\beta}_1) \dots (\lambda - \bar{\beta}_M) &= \bar{B}(\lambda). \end{aligned}$$

Remarquons que, pour que $\varphi_s(\lambda)$ vérifie (3.17), il est nécessaire que $\Phi^{(1)}(\lambda)$ et $\Phi^{(2)}(\lambda)$ aient mêmes pôles.

L'ensemble de ceux-ci, pour que (3.22) soit vérifié, doit comprendre les pôles (avec les mêmes ordres de multiplicité) de $M^{(1)}(\lambda) \frac{1}{(\lambda - \gamma_1) \dots (\lambda - \gamma_{L_1})}$ et peut comprendre les racines (avec les mêmes ordres de multiplicité) de $B(\lambda) \bar{B}(\lambda)$.

Écrivons donc

$$(3.25) \quad \Phi^{(1)}(\lambda) = \frac{\omega^{(1)}(\lambda)}{(\lambda - \delta_1) \dots (\lambda - \delta_\rho) (\lambda - \gamma_1) \dots (\lambda - \gamma_{L_1}) B(\lambda) \bar{B}(\lambda)},$$

$$(3.26) \quad \Phi^{(2)}(\lambda) = \frac{\omega^{(2)}(\lambda)}{(\lambda - \delta_1) \dots (\lambda - \delta_\rho) (\lambda - \gamma_1) \dots (\lambda - \gamma_{L_1}) B(\lambda) \bar{B}(\lambda)}.$$

Pour que $\Phi^{(1)}(\lambda)$ et $\Phi^{(2)}(\lambda)$ appartiennent à \mathcal{L}_F^2 il faut et il suffit que les degrés des polynômes $\omega^{(1)}(\lambda)$ et $\omega^{(2)}(\lambda)$ soit inférieurs ou égaux à $p + M + N + L_1 - 1$.

Nous allons voir que nous pouvons trouver de tels polynômes $\omega^{(1)}(\lambda)$ et $\omega^{(2)}(\lambda)$ qui vérifient (3.17), (3.21) et (3.22).

Pour que (3.17) soit satisfait complètement il faut et il suffit que $\Phi^{(1)}(\lambda) + e^{-i\lambda T} \Phi^{(2)}(\lambda)$ soit une fonction entière; ce qui fournit $p + 2M + L_1$ équations; (3.21) impose n équations.

Passons à la vérification de (3.22) :

$\omega^{(1)}(\lambda)$ et $\omega^{(2)}(\lambda)$ doivent avoir mêmes zéros que respectivement $A(\lambda)$ et $\bar{A}(\lambda)$.

$$e^{is\lambda} D(\lambda) A(\lambda) \bar{A}(\lambda) (\lambda - \delta_1) \dots (\lambda - \delta_\rho) - \omega^{(1)}(\lambda) (\lambda - \gamma_1) \dots (\lambda - \gamma_{L_1})$$

doit avoir les mêmes zéros que $(\lambda - \gamma_1) \dots (\lambda - \gamma_{L_1})$.

Enfin $\psi_1(\lambda)$ doit rester finie pour les δ_i dont la partie imaginaire est positive ou nulle et $\psi_2(\lambda)$ doit rester finie pour les δ_i dont la partie imaginaire est négative ou nulle.

Nous remarquerons que ces dernières conditions ne fournissent que p équations; en effet pour λ réel $\psi_1(\lambda) + e^{-i\lambda T} \psi_2(\lambda)$ est toujours fini, donc $\psi_1(\lambda)$ finie entraîne $\psi_2(\lambda)$ finie quand λ réel.

Au total si nous prenons $\omega^{(1)}(\lambda)$ et $\omega^{(2)}(\lambda)$ des polynômes de degré $p + M + N + L_1 - 1$; les $2p + 2M + 2N + 2L_1 + n$ inconnues constituées par les coefficients des polynômes $\omega^{(1)}(\lambda)$ et $\omega^{(2)}(\lambda)$ et les coefficients c_1, \dots, c_n doivent vérifier

$$p + 2M + L_1 + n + 2N + L_1 + p = 2p + n + 2M + 2N + 2L_1$$

équations linéaires par rapport à ces inconnues.

On peut montrer que le déterminant du système est différent de zéro.

Par suite, il existe une fonction $\varphi_s(\lambda)$ de la forme (3.17) qui vérifie (a), (b) et (c).

On a alors

$$\tilde{Z} = \int_{-\infty}^{+\infty} \varphi_s(\lambda) d\xi(\lambda).$$

L'unicité d'une fonction $\varphi_s(\lambda)$ vérifiant (a), (b) et (c) résulte de celle de \tilde{Z} et de l'isométrie entre \mathcal{H}_{X_0} et $\mathcal{L}_{\mathbb{F}}^2$.

Si nous écrivons, comme nous l'avons indiqué plus haut [(3.18) et (3.19)]

$$\Phi^{(1)}(\lambda) + e^{-i\lambda T} \Phi^{(2)}(\lambda) = \sum_{k=1}^{N-M-1} (\omega_k + \omega'_k e^{-i\lambda T}) (i\lambda)^k + \int_{-T}^0 e^{i\lambda t} \nu(t) dt,$$

alors

$$\tilde{Z} = \sum_{k=1}^{N-M-1} [\omega_k X_0^{(k)}(0) + \omega'_k X_0^{(k)}(-T)] + \int_{-T}^0 X_0(t) \nu(t) dt + \sum_{j=1}^n P_j X_0(-t_j);$$

et, d'après (3.21),

$$\tilde{Z} = \sum_{k=1}^{N-M-1} [\omega_k X^{(k)}(0) + \omega'_k X^{(k)}(-T)] + \int_{-T}^0 X(t) \nu(t) dt + \sum_{j=1}^n P_j X(-t_j)$$

et

$$(3.27) \quad \tilde{Y}(s) = \sum_{k=1}^{N-M-1} [\omega_k X^{(k)}(0) + \omega'_k X^{(k)}(-T)] + \int_{-T}^0 X(t) \nu(t) dt.$$

Calcul de l'erreur

$$E | Y(s) - \tilde{Y}(s) |^2 = E [Z - \tilde{Z}] [\overline{Z - \tilde{Z}}].$$

Mais nous avons vu dans la démonstration du théorème 7 qu'une conséquence de l'équation (3.8) était que

$$E [Z - \tilde{Z}] \bar{\tilde{Z}} = 0.$$

Donc,

$$\begin{aligned} E | Y(s) - \tilde{Y}(s) |^2 &= E [Z - \tilde{Z}] \bar{Z} \\ &= E [Z - \tilde{Z}] \left[Y(s) + \sum_{j=1}^n P_j \bar{X}(-t_j) \right] \\ &= E [Z - \tilde{Z}] \bar{Y}_0(s) + \sum_{j=1}^n P_j \sum_{k=1}^n c_k g_k(-t_j) \\ &= E \left\{ Y_0(s) - \int \left(\varphi_s(\lambda) - \sum_{j=1}^n P_j e^{-i\lambda t_j} \right) d\xi(\lambda) \right\} Y_0(s) \\ &\quad - \sum_{k=1}^n c_k g_k(s), \end{aligned}$$

$$(3.28) \quad \mathbf{E} | Y(s) - \tilde{Y}(s) |^2 = \mathbf{E} | Y_0(s) |^2 - \int_{-\infty}^{+\infty} \left[\varphi_s(\lambda) - \sum_{j=1}^n P_j e^{-i\lambda t_j} \right] f_1(\lambda) d\lambda \\ - \sum_{k=1}^n c_k g_k(s).$$

BIBLIOGRAPHIE.

- [A₁] B. P. ADHIKARI, *Tests d'hypothèses pour processus stochastiques* (Thèse de doctorat, Paris, 1957).
- [B₁] P. BÉTHOUX, *Discrimination entre plusieurs signaux en télécommunication* (C. R. Acad. Sc., t. 247, 1958, p. 412). *Nombre maximal de signaux d'énergie totale fixée parmi lesquels on peut discriminer à ε près en présence d'un bruit blanc Gaussien* (C. R. Acad. Sc., t. 247, 1958, p. 573). *Filtrage d'une fonction aléatoire dont la moyenne est une fonction linéaire* (C. R. Acad. Sc., t. 248, 1959, p. 3685).
- [B₂] A. BLANC-LAPIERRE et R. FORTET, *Théorie des fonctions aléatoires*, Masson, Paris, 1953.
- [C₁] CHIANG TSE PEI, *On the estimation of regression coefficients of a continuous parameter times series with stationary residual* [Théorie des probabilités et ses applications (Journal russe, t. IV, n° 4, 1959, p. 405)].
- [D₁] J. L. DOOB, *Stochastic processes*, New York, 1953.
- [D₂] N. DUNFORD et J. SCHWARTZ, *Linear operators*, New York, 1958.
- [F₁] R. FORTET, *Hypothesis testing and estimation for random laplacian functions* (Fourth Berkeley symposium on Mathematical Statistics and Probability).
- [F₂] R. FORTET et E. MOURIER, *Les fonctions aléatoires comme éléments aléatoires dans un espace de Banach* (J. Math. Pures et Appl., t. 38, 1959, p. 347-364).
- [G₁] U. GRENANDER, *Stochastic processes and statistical inference* (Arkiv för Math., t. 1, 1950, p. 195-279).
- [H₁] P. HALMOS, *Measure theory*, New York, 1950.
- [H₂] P. HALMOS et L. SAVAGE, *Applications of the Radon-Nikodym theorem to the theory of sufficient statistics* (Ann. Math. Stat., t. 20, 1959, p. 225).
- [H₃] A. HANEN, *Propriétés du maximum de vraisemblance pour des processus laplaciens* (C. R. Acad. Sc., t. 250, 1960, p. 4268).
- [H₄] A. HANEN, *Quelques problèmes de tests d'hypothèses et d'estimation en théorie des communications* (C. R. Acad. Sc., t. 250, 1960, p. 3940).
- [J₁] M. JIRINA, *Probabilités conditionnelles sur des σ -algèbres strictement séparables* (J. Math. Tch., t. 4, 1954, p. 372-380).
- [K₁] G. KALLIANPUR, *A problem in optimum filtering with finite data* (Ann. Math. Stat., t. 30, 1959, p. 659).
- [L₁] P. LÉVY, *Processus stochastiques et mouvement brownien*, Gauthier-Villars, Paris, 1948.
- [M₁] H. B. MANN, *A theory of estimation for fundamental random process and Ornstein-Uhlenbeck process* (Sankhya t. 13, part 4, 1954 p. 325-350).
- [M₂] E. MOURIER, *Éléments aléatoires à valeurs dans un espace de Banach* (Ann. Inst. H. Poincaré, t. 13, 1953, p. 161-244).
- [P₁] E. PARSEN, *Regression analysis of continuous parameter time series* Department Statist., Stanford Univ., Technical Report n° 33, Juin 1960.
- [S₁] C. SHANNON et W. WEAVER, *The mathematical theory of communication*, Urbana, Univ. Illinois Press, 1949.

- [S₂] C. STRIEBEL, *Densities for stochastic processes* (*Ann. Math. Stat.*, t. 30, n° 2, 1959, p. 559).
- [W₁] N. WIENER et P. MASANI, *Non linear prediction* (*Probability and Statistics*, The H. Cramér Vol., Stockholm).
- [Y₁] A. M. YAGLOM, *Extrapolation, interpolation, filtrage des processus aléatoires stationnaires à densité spectrale rationnelle* (*Trudy Moskov, Mat.*, Obsc. 4, 1955, p. 333-374).
- [Y₂] A. M. YAGLOM, *Théorie de la corrélation des processus à accroissements n^{ièmes} stationnaires* (*Recueil Math. Acad. Sc.*, Moscou, t. 37, (79), 1956, p. 141).
- [Y₃] A. M. YAGLOM, *Solutions effectives de problèmes d'approximations linéaires pour des processus vectoriels stationnaires à spectres rationnels* [*Théorie des probabilités et ses applications* (Journal russe), t. V, n° 3, 1960].
-

TABLE DES MATIÈRES.

	Pages.
SOMMAIRE.....	255
INTRODUCTION.....	255
CHAPITRE I. — <i>Tests sur la moyenne d'une fonction aléatoire :</i>	
Application au problème de la vitesse de transmission de l'information à travers un canal de communication de bande passante finie.	
1. Test sur la moyenne d'une fonction aléatoire.....	259
2. Cas stationnaire. Canal de communication de bande passante finie.....	263
3. Discrimination entre deux signaux.....	266
4. Discrimination entre plus de deux signaux.....	270
5. Capacité d'un canal de communication perturbé par un bruit blanc laplacien.	274
CHAPITRE II. — <i>Estimation par le maximum de vraisemblance de la moyenne d'une fonction aléatoire permanente.</i>	
1. Position du problème et notations.....	284
2. Estimateur du maximum de vraisemblance.....	285
3. Cas où \mathcal{F} est un sous-espace vectoriel de dimension finie de H	290
4. Quelques autres cas.....	296
CHAPITRE III. — <i>Certains problèmes de prévision.</i>	
1. Prévision à partir de l'observation sur un intervalle d'une fonction aléatoire permanente.....	300
2. Un problème de détermination d'un filtrage optimal.....	304
3. Prévision linéaire dans le cas des fonctions aléatoires dont les moyennes sont des combinaisons linéaires inconnues de fonctions certaines connues.	308
BIBLIOGRAPHIE.....	320
